

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 09:55:43 ; Search time 45 Seconds
(without alignments)
2475.065 Million cell updates/sec

Title: US-09-466-778B-11
Perfect score: 353
Sequence: 1 MTGPGHKCECKSHYVDGL.....ALAAYSFRINRKTIGXFHF 353

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1017041 seqs, 315518202 residues

Word size: 9
Total number of hits satisfying chosen parameters: 12

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 500 summaries

Database : SPTREMBL 25: *
1: sp_archaea: *
2: sp_bacteria: *
3: sp_fungi: *
4: sp_human: *
5: sp_invertebrate: *
6: sp_mammal: *
7: sp_mhc: *
8: sp_organelle: *
9: sp_phage: *
10: sp_plant: *
11: sp_rodent: *
12: sp_virus: *
13: sp Vertebrate: *
14: sp Unclassified: *
15: sp_virus: *
16: sp_bacteriap: *
17: sp_archaeap: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	98	27.8	1192	4 Q9H7H7	Q9h7h7 homo sapien
2	98	27.8	1416	4 Q86UR4	Q86ur4 homo sapien
3	98	27.8	1736	4 Q8RES1	Q8ces1 homo sapien
4	98	27.8	2551	4 Q8WQ08	Q8wq08 homo sapien
5	98	27.8	2551	4 Q8IUG9	Q8iug9 homo sapien
6	98	27.8	2551	4 Q7Z5N9	Q7z5n9 homo sapien
7	93	26.3	897	4 Q9NRV3	Q9nrv3 homo sapien
8	93	26.3	1069	4 Q9UF98	Q9uf98 homo sapien
9	32	9.1	500	6 Q864U4	Q864u4 bos taurus
10	30	8.5	894	11 Q8BM87	Q8bm87 mus musculu
11	30	8.5	2559	11 Q8R4U0	Q8r4u0 mus musculu
12	25	7.1	1431	11 Q8CFM6	Q8cfm6 rattus norv

ALIGNMENTS

RESULT 1
Q9H7H7 PRELIMINARY; PRT; 1192 AA.
AC Q9H7H7;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
GN FLJ00112 protein (Fragment).
GN FLJ00112;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue=Spleen;
RA Ohara O., Nagase T., Kikuno R., Okumura K.;
RT "The nucleotide sequence of a long cDNA clone isolated from human spleen."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK024503; BAB15793.1; -
DR HSP; P98066; ITSG.
DR GO; GO:0005540; F:hyaluronic acid binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR000782; BIGH3 FAS1.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR000538; Link.
DR Pfam; PF00008; EGF; 7.
DR Pfam; PF02469; Fasciclin; 2.
DR Pfam; PF00193; Xlink; 1.
DR PRINTS; PR00011; EGF_LAMININ.
DR PRINTS; PR01265; LINKMODULE.
DR PRODOM; PD000918; Link; 1.
DR SMART; SM00181; EGF; 9.
DR SMART; SM00445; LINK; 1.
DR SMART; SM00554; FAS1; 3.
DR PROSITE; PS00022; EGF_1; 4.
DR PROSITE; PS01186; EGF_2; 8.
DR PROSITE; PS02113; FAS1; 3.
DR PROSITE; PS01248; LAMININ TYPE EGF; 2.
KW EGF-like domain; Laminin EGF-like domain.
FT NON TER 1
SQ SEQUENCE 1192 AA; 128738 MW; 5966BED4B83BD9C1 CRC64;

Query Match 27.8%; Score 98; DB 4; Length 1192;
Best Local Similarity 100.0%; Pred. No. 1.4e-94;
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 169 SGNLQVIMSPSLTNFLTEVLAYSNSSARGRAFLHLTDLISIRGTLFVPNSGLGENET 228
Db 952 SGNLQVIMSPSLTNFLTEVLAYSNSSARGRAFLHLTDLISIRGTLFVPNSGLGENET 1011
Qy 229 LSGRDIHHLANVSMFFYNDLVNGTTLQTRLGSKLIT 266
Db 1012 LSGRDIHHLANVSMFFYNDLVNGTTLQTRLGSKLIT 1049

RESULT 2

Q86UR4 PRELIMINARY; PRT; 1416 AA.
AC Q86UR4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hyaluronan receptor for endocytosis (Fragment).
GN HARE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22588541; PubMed=12626425;
RA Zhou B., McGary C.T., Weigel J.A., Saxena A., Weigel P.H.;
RT "Purification and molecular identification of the human hyaluronan
RT receptor for endocytosis.";
RL Glycobiology 13:339-349 (2003).
DR EMBL; AY227444; AAO39681.1; -
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005540; F:hyaluronic acid binding; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0005198; P:cell adhesion; IEA.
DR GO; GO:0005198; P:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR000782; BIGH3_FAS1.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR00538; Link.
DR InterPro; IPR001981; EGF Ca.
DR InterPro; IPR006210; EGF-like.
DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR00538; Link.
DR Pfam; PF00008; EGF; 8.
DR Pfam; PF02469; Fasciclin; 3.
DR Pfam; PF00193; Xlink; 1.
DR PRINTS; PR00011; EGF_LAMININ.
DR ProDom; PD000918; Link; 1.
DR SMART; SM00181; EGF; 11.
DR SMART; SM00179; EGF CA; 4.
DR SMART; SM00554; FAS1; 4.
DR SMART; SM00445; LINK; 1.
DR PROSITE; PS00022; EGF 1; 4.
DR PROSITE; PS01186; EGF 2; 8.
DR PROSITE; PS02133; FAS1; 4.
DR PROSITE; PS01248; LAMININ_TYPE_EGF; 2.
KW Receptor.
FT NON_TER 1
FT CHAIN <1 1416
FT SEQUENCE 1416 AA; 154089 MW; 85A216D38E3B10DE CRC64;
Query Match 27.8%; Score 98; DB 4; Length 1416;
Best Local Similarity 100.0%; Pred. No. 1.6e-94;
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 169 SGNLLQVMSFPSTLNFTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENET 228
Db 1176 SGNLLQVMSFPSTLNFTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENET 1235
QY 229 LSGRDIEHHLANVSMFFYNDLVNGTTLQTLRGLSKLLIT 266
Db 1236 LSGRDIEHHLANVSMFFYNDLVNGTTLQTLRGLSKLLIT 1273
Query Match 27.8%; Score 98; DB 4; Length 1416;
Best Local Similarity 100.0%; Pred. No. 1.6e-94;
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 169 SGNLLQVMSFPSTLNFTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENET 228
Db 1176 SGNLLQVMSFPSTLNFTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENET 1235
QY 229 LSGRDIEHHLANVSMFFYNDLVNGTTLQTLRGLSKLLIT 266
Db 1236 LSGRDIEHHLANVSMFFYNDLVNGTTLQTLRGLSKLLIT 1273
RESULT 3
Q8TES1
ID Q8TES1 PRELIMINARY; PRT; 1736 AA.
AC Q8TES1;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE FLJ00122 protein (Fragment).
GN FLJ00122.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Spleen;
RA Jikuya H., Takano J., Nomura N., Kikuno R., Nagase T., Ohara O.;
RT "The nucleotide sequence of a long cDNA clone isolated from human
RT spleen.";
```

```
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK074051; BAB84877.1; -
DR GO; GO:0005540; F:hyaluronic acid binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR000782; BIGH3_FAS1.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR000538; Link.
DR Pfam; PF00008; EGF; 12.
DR Pfam; PF02469; Fasciclin; 4.
DR Pfam; PF00193; Xlink; 1.
DR PRINTS; PR00011; EGF_LAMININ.
DR ProDom; PD000918; Link; 1.
DR SMART; SM00181; EGF; 15.
DR SMART; SM00554; FAS1; 5.
DR SMART; SM00445; LINK; 1.
DR PROSITE; PS00022; EGF 1; 4.
DR PROSITE; PS01186; EGF 2; 12.
DR PROSITE; PS02133; FAS1; 5.
DR PROSITE; PS01248; LAMININ_TYPE_EGF; 2.
KW EGF-like domain.
FT NON_TER 1
FT SEQUENCE 1736 AA; 187887 MW; 50982047B43925F2 CRC64;
Query Match 27.8%; Score 98; DB 4; Length 1736;
Best Local Similarity 100.0%; Pred. No. 1.9e-94;
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 169 SGNLLQVMSFPSTLNFTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENET 228
Db 1496 SGNLLQVMSFPSTLNFTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENET 1555
QY 229 LSGRDIEHHLANVSMFFYNDLVNGTTLQTLRGLSKLLIT 266
Db 1556 LSGRDIEHHLANVSMFFYNDLVNGTTLQTLRGLSKLLIT 1593
RESULT 4
Q8WQ08
ID Q8WQ08 PRELIMINARY; PRT; 2551 AA.
AC Q8WQ08;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Stabilin-2.
GN STAB2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Politz O., Gratchev A., McCourt P.A.G., Schledzewski K., Guillot P.,
RA Johansson S., Birk R., Hakiy N., Franke P., Kodelja V., Kannicht C.,
RA Orfanos C.E., Johansson S., Goerdts S.;
RT "Stabilin-1 and stabilin-2 constitute a novel family of fasciclin
RT domain-containing adhesion molecules associated with endothelial-
RT macrophage differentiation and angiogenic processes.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ295695; CAC82105.1; -
DR GO; GO:0005540; F:hyaluronic acid binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR000782; BIGH3_FAS1.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR00538; Link.
DR Pfam; PF00008; EGF; 16.
DR Pfam; PF02469; Fasciclin; 5.
DR Pfam; PF00193; Xlink; 1.
```

```
DR PRINTS; PRO00011; EGFLAMININ.
DR PRODOM; PD000918; Link; 1.
DR SMART; SM00181; EGF; 20.
DR SMART; SM00554; FAS1; 7.
DR SMART; SM00445; LINK; 1.
DR PROSITE; PS00022; EGF_1; 7.
DR PROSITE; PS01186; EGF_2; 16.
DR PROSITE; PS02013; FAS1; 7.
DR PROSITE; PS01248; LAMININ TYPE EGF; 2.
DR PROSITE; PS01248; LAMININ EGF-like domain.
KW EGF-like domain; Laminin EGF-like domain.
SQ SEQUENCE 2551 AA; 276992 MW; 60A44651CCC2B69 CRC64;

Query Match 27.8%; Score 98; DB 4; Length 2551;
Best Local Similarity 100.0%; Pred. No. 2.8e-94;
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 169 SGNLLQVLMSFSLTNFLTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENET 228
Db 2311 SGNLLQVLMSFSLTNFLTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENET 2370

Qy 229 LSGRDIHHLANVMFFYNDLVNGTTLQTRLGSKLLIT 266
Db 2371 LSGRDIHHLANVMFFYNDLVNGTTLQTRLGSKLLIT 2408

RESULT 5
Q8IU99 PRELIMINARY; PRT; 2551 AA.
AC Q8IU99;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE FELE-2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Adachi H., Tsujimoto M.;
RT "FEEL-1, a novel scavenger receptor with in vitro bacteria-binding and
RT angiogenesis-modulating activities.";
RL J. Biol. Chem. 277:34264-34270 (2002).
DR EMBL; AB052958; BAC15608.1; -;
DR Genbank; HGNC:18629; STAB2.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005540; F:hyaluronic acid binding; IEA.
DR GO; GO:0004867; F:serine protease inhibitor activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR000782; BIGH3_FAS1.
DR InterPro; IPR001881; EGF_CA.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; LEGF.
DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR000538; Link.
DR InterPro; IPR000737; Squash.
DR Pfam; PF00008; EGF; 16.
DR Pfam; PF02469; Fasciclin; 6.
DR Pfam; PF00193; Xlink; 1.
DR PRINTS; PRO0011; EGFLAMININ.
DR PRODOM; PD000918; Link; 1.
DR SMART; SM00181; EGF; 22.
DR SMART; SM00179; EGF_CA; 8.
DR SMART; SM00180; EGF_Lam; 5.
DR SMART; SM00554; FAS1; 7.
DR SMART; SM00445; LINK; 1.
DR SMART; SM00286; PTI; 8.
DR PROSITE; PS00022; EGF_1; 7.
DR PROSITE; PS01186; EGF_2; 16.
DR PROSITE; PS02013; FAS1; 7.
```

```
DR PROSITE; PS01248; LAMININ TYPE EGF; 2.
DR SQ SEQUENCE 2551 AA; 277026 MW; 3ADDF3F5BD2019C4 CRC64;

Query Match 27.8%; Score 98; DB 4; Length 2551;
Best Local Similarity 100.0%; Pred. No. 2.8e-94;
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 169 SGNLLQVLMSFSLTNFLTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENET 228
Db 2311 SGNLLQVLMSFSLTNFLTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENET 2370

Qy 229 LSGRDIHHLANVMFFYNDLVNGTTLQTRLGSKLLIT 266
Db 2371 LSGRDIHHLANVMFFYNDLVNGTTLQTRLGSKLLIT 2408

RESULT 6
Q7Z5N9 PRELIMINARY; PRT; 2551 AA.
AC Q7Z5N9;
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE FEX2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Park S.-Y., Kim I.-S.;
RT "FEX2, a novel cell adhesion molecule of Fas-1 superfamily mediates
RT cell-cell interaction.";
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY311388; AAF74958.1; -;
DR SQ SEQUENCE 2551 AA; 276986 MW; 3ACB6A6C3CB80044 CRC64;

Query Match 27.8%; Score 98; DB 4; Length 2551;
Best Local Similarity 100.0%; Pred. No. 2.8e-94;
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 169 SGNLLQVLMSFSLTNFLTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENET 228
Db 2311 SGNLLQVLMSFSLTNFLTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENET 2370

Qy 229 LSGRDIHHLANVMFFYNDLVNGTTLQTRLGSKLLIT 266
Db 2371 LSGRDIHHLANVMFFYNDLVNGTTLQTRLGSKLLIT 2408

RESULT 7
Q9NR93 PRELIMINARY; PRT; 897 AA.
AC Q9NR93;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE CD44-like precursor FELL.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tao Q., Zhang W., Cao X.;
RT "Molecular cloning and characterization of human FELL sharing homology
RT with CD44.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF160476; AAF82398.1; -;
DR HSP; P98066; ITSG.
DR GO; GO:0005540; F:hyaluronic acid binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
```

DR InterPro: IPR000782; BIGH3_FAS1.
DR InterPro: IPR006209; EGF-like.
DR InterPro: IPR006210; IEGF.
DR InterPro: IPR002049; Laminin_EGF.
DR InterPro: IPR000538; Link.
DR Pfam: PF00008; EGF; 4.
DR Pfam: PF02469; Fasciclin; 2.
DR Pfam: PF00193; Xlink; 1.
DR PRINTS: PR01265; LINKMODULE.
DR ProDom: PD000918; Link; 1.
DR SMART: SM00181; EGF; 5.
DR SMART: SM00554; FAS1; 3.
DR SMART: SM00445; LINK; 1.
DR PROSITE: PS00022; EGF 1; 2.
DR PROSITE: PS01186; EGF 2; 4.
DR PROSITE: PS02113; FAS1; 3.
DR PROSITE: PS01248; LAMININ_TYPE_EGF; 1.
KW EGF-like domain.
SQ SEQUENCE 897 AA; 97585 MW; EB920AF36101E388 CRC64;
Query Match 26.3%; Score 93; DB 4; Length 897;
Best Local Similarity 100.0%; Pred. No. 2.2e-89;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MTGPGKHCKCKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVGFHL 60
Db 489 MTGPGKHCKCKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVGFHL 548
Qy 61 RSLPGQYKLTDFDKAREACANEATMATYNQLSY 93
Db 549 RSLPGQYKLTDFDKAREACANEATMATYNQLSY 581
RESULT 8
Q9UF98 PRELIMINARY; PRT; 1069 AA.
ID Q9UF98
AC Q9UF98
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE Hypothetical protein (fragment).
GN DKF2P434E0321.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Blum H., Bauersachs S., Mewes H.W., Gassenhuber J., Wiemann S.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL133021; CAB61358.1; -
DR FIR; T42681; T42681.
DR HSSP; P98066; ITSG.
DR GO; GO:0005540; F-hyaluronic acid binding; IEA.
DR GO; GO:0005198; F-structural molecule activity; IEA.
DR GO; GO:0007155; P-cell adhesion; IEA.
DR InterPro: IPR000782; BIGH3_FAS1.
DR InterPro: IPR006209; EGF-like.
DR InterPro: IPR006210; IEGF.
DR InterPro: IPR002049; Laminin_EGF.
DR Pfam: PF00008; EGF; 7.
DR Pfam: PF02469; Fasciclin; 2.
DR PRINTS: PR00193; Xlink; 1.
DR PRINTS: PR00011; EGF_LAMININ.
DR PRINTS: PR01265; LINKMODULE.
DR ProDom: PD000918; Link; 1.
DR SMART: SM00181; EGF; 8.
DR SMART: SM00554; FAS1; 3.
DR SMART: SM00445; LINK; 1.
DR PROSITE: PS00022; EGF 1; 2.
DR PROSITE: PS01186; EGF 2; 5.

DR PROSITE: PS02113; FAS1; 3.
DR PROSITE: PS01248; LAMININ_TYPE_EGF; 1.
KW Hypothetical protein; EGF-like domain; Laminin EGF-like domain.
FT NON_TER 1
SQ SEQUENCE 1069 AA; 115666 MW; 04B0960932164D63 CRC64;
Query Match 26.3%; Score 93; DB 4; Length 1069;
Best Local Similarity 100.0%; Pred. No. 2.5e-99;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MTGPGKHCKCKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVGFHL 60
Db 697 MTGPGKHCKCKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVGFHL 756
Qy 61 RSLPGQYKLTDFDKAREACANEATMATYNQLSY 93
Db 757 RSLPGQYKLTDFDKAREACANEATMATYNQLSY 789
RESULT 9
Q864U4 PRELIMINARY; PRT; 500 AA.
ID Q864U4
AC Q864U4
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative hyaluronan receptor for endocytosis (fragment).
GN HARE.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Ulbrich S.E., Thoenes S.;
RL "Expression of the hyaluronan system in the bovine oviduct.";
RN Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Einspanier R.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ550060; CAD79334.2; -
KW Receptor.
FT NON_TER 1
FT NON_TER 500
SQ SEQUENCE 500 AA; 54479 MW; C93AA6404B103582 CRC64;
Query Match 9.1%; Score 32; DB 6; Length 500;
Best Local Similarity 100.0%; Pred. No. 8.2e-25;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 48 LHFQDTTGVGFHLRSLPGQYKLTDFDKAREACA 79
Db 422 LHFQDTTGVGFHLRSLPGQYKLTDFDKAREACA 453
RESULT 10
Q8BM87 PRELIMINARY; PRT; 894 AA.
ID Q8BM87
AC Q8BM87
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to CD44-like precursor FELL (fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.


```
RC STRAIN=C57BL/6J; TISSUE=Dienecephalon;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60, 770 full-length cDNAs";
RL Nature 420:563-573 (2002).
DR EMBL; AK034522; BAC28741.1; -.
DR GO; GO:0005540; F:hyaluronic acid binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR000782; BIGH3_FAS1.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR000538; Link.
DR Pfam; PF00008; EGF; 3.
DR Pfam; PF02469; Fasciclin; 2.
DR PRINTS; PR01265; LINKMODULE.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00554; FAS1; 3.
DR SMART; SM00445; LINK; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 4.
DR PROSITE; PS0213; FAS1; 3.
DR PROSITE; PS01248; LAMININ_TYPE_EGF; 1.
FT NON TER 1
SQ SEQUENCE 894 AA; 97059 MW; 3286E2A223CFBB96 CRC64;

Query Match 8.5%; Score 30; DB 11; Length 894;
Best Local Similarity 100.0%; Pred. No. 1.9e-22;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 CVDLHFQDTTGVFHLRSLPGQYKLTFDKA 74
Db 530 CVDLHFQDTTGVFHLRSLPGQYKLTFDKA 559

RESULT 11
Q8R4U0 PRELIMINARY; PRT; 2559 AA.
AC Q8R4U0;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Stabilin-2.
GN STAB2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Liver;
RX MEDLINE=21818465; PubMed=11829752;
RA Politz O., Gratchev A., McCourt P.A.G., Schledzewski K., Guillot P.,
RA Johansson S., Svineng F.W., Franke P., Kannicht C., Kzyshkowska J.,
RA Longati P., Velten G., Johansson S., Goerdts S.,
RT "Stabilin-1 and -2 constitute a novel family of fasciclin-like
RT hyaluronan receptor homologues";
RL Biochem. J. 362:155-164 (2002).
DR EMBL; AF364951; AAL91684.2; -.
DR MGD; MGI:2178743; Stab2.
DR GO; GO:0003779; F:actin binding; IEA.
DR GO; GO:0005540; F:hyaluronic acid binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR001589; Actbind actnin.
DR InterPro; IPR000782; BIGH3_FAS1.
DR InterPro; IPR006209; EGF_like.
```

```
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR002049; Laminin_EGF.
DR Pfam; PF000538; Link.
DR Pfam; PF00008; EGF; 13.
DR Pfam; PF02469; Fasciclin; 6.
DR Pfam; PF00193; Xlink; 1.
DR PRINTS; PR00011; EGF_LAMININ.
DR ProDom; PD000918; Link; 1.
DR SMART; SM00181; EGF; 21.
DR SMART; SM00554; FAS1; 7.
DR SMART; SM00445; LINK; 1.
DR PROSITE; PS00019; ACTININ 1; 1.
DR PROSITE; PS00022; EGF_1; 7.
DR PROSITE; PS01186; EGF_2; 16.
DR PROSITE; PS0213; FAS1; 7.
DR PROSITE; PS01248; LAMININ_TYPE_EGF; 2.
KW EGF-like domain.
SQ SEQUENCE 2559 AA; 277530 MW; 1C9855AD61EFF015 CRC64;

Query Match 8.5%; Score 30; DB 11; Length 2559;
Best Local Similarity 100.0%; Pred. No. 5e-22;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 CVDLHFQDTTGVFHLRSLPGQYKLTFDKA 74
Db 2195 CVDLHFQDTTGVFHLRSLPGQYKLTFDKA 2224

RESULT 12
Q8CFM6 PRELIMINARY; PRT; 1431 AA.
AC Q8CFM6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hyaluronan receptor for endocytosis HARE precursor (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=22169209; PubMed=12181351;
RA Zhou B., Weigel J.A., Saxena A., Weigel P.H.;
RT "Molecular Cloning and Functional Expression of the Rat 175-kDa
RT Hyaluronan Receptor for Endocytosis.";
RL Mol. Biol. Cell 13:2853-2868 (2002).
DR EMBL; AY007370; AAG13634.1; -.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005540; F:hyaluronic acid binding; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR000782; BIGH3_FAS1.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR000538; Link.
DR Pfam; PF00008; EGF; 9.
DR Pfam; PF02469; Fasciclin; 3.
DR Pfam; PF00193; Xlink; 1.
DR PRINTS; PR01265; LINKMODULE.
DR ProDom; PD000918; Link; 1.
DR SMART; SM00181; EGF; 11.
DR SMART; SM00179; EGF_Ca; 2.
DR SMART; SM00554; FAS1; 4.
DR SMART; SM00445; LINK; 1.
DR PROSITE; PS00022; EGF_1; 3.
DR PROSITE; PS01186; EGF_2; 8.
DR PROSITE; PS0213; FAS1; 4.
DR PROSITE; PS01248; LAMININ_TYPE_EGF; 2.
```

KW Receptor.
 PT NON_TER 1
 SQ SEQUENCE 1431 AA; 155914 MW; 3340A8F7AF366722 CRC64;
 Query Match 7.1%; Score 25; DB 11; Length 1431;
 Best Local Similarity 100.0%; Pred. No. 5.9e-17;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 50 FQDTTVGVFHLRSPLGQYKLTFDKA 74
 Db 1057 FQDTTVGVFHLRSPLGQYKLTFDKA 1081

Search completed: April 1, 2004, 09:58:24
 Job time : 46 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 09:55:43 ; Search time 17 seconds
(without alignments)
1081.221 Million cell updates/sec

Title: US-09-466-778B-11
Perfect score: 353
Sequence: 1 MTGPGKHCKCKSHYVGDGL.....ALAAYSYRINRKTIGFXHF 353

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 141681 seqs, 52070155 residues

Word size: 9

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 500 summaries

Database: SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description

No matches found

Search completed: April 1, 2004, 09:57:25
Job time : 17 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 09:55:43 ; Search time 20 Seconds
(without alignments)
1697.780 Million cell updates/sec

Title: US-09-466-778B-11

Perfect score: 353

Sequence: 1 MTGPKHKCKSHYVGDGL.....ALAAYSYFRINRKTIGEXHF 353

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283366 seqs, 96191526 residues

Word size : 9

Total number of hits satisfying chosen parameters: 1

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 500 summaries

Database :

PIR 78.*
1: Pirl.*
2: Pirl2.*
3: Pirl3.*
4: Pirl4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	93	26.3	1069	2 T42681	hypothetical prote

ALIGNMENTS

RESULT 1

T42681

hypothetical protein DKFZp434E0321.1 - human (fragments)

C;Species: Homo sapiens (man)

C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 11-Jan-2000

C;Accession: T42681

R;Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, November 1999

A;Reference number: Z22233

A;Accession: T42681

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-870;871-1069 <AAA>

A;Cross-references: EMBL:AL13021

A;Experimental source: adult testis; clone DKFZp434E0321

A;Note: the cDNA sequence contains a +1 frameshift near codon 870

C;Genetics:

A;Note: DKFZp434E0321.1

Query Match

Best Local Similarity 100.0%; Pred. No. 6.6e-88; Length 1069;

Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTGPKHKCKSHYVGDGLNCEPEQLPIDRCLQDNGOCHADAKCVDLHFQDITTVGVFHL 60
Db 597 MTGPKHKCKSHYVGDGLNCEPEQLPIDRCLQDNGOCHADAKCVDLHFQDITTVGVFHL 756

Qy 61 RSPGLQYKLTDFDKAREACANEAAATMATYNQLSY 93

Db 757 RSPGLQYKLTDFDKAREACANEAAATMATYNQLSY 789

Search completed: April 1, 2004, 09:58:56
Job time : 20 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 09:58:28 ; Search time 45 Seconds
(without alignments)
2057.758 Million cell updates/sec

Title: US-09-466-778B-11

Perfect score: 353
Sequence: 1 MTGPKHCKCKSHYVDGL.....ALAAYSYFRINRKTIGFHF 353

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1069545 seqs, 262320428 residues

Word size: 9

Total number of hits satisfying chosen parameters: 23

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 500 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubaa/US09D_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length DB	ID	Description
1	98	27.8	1192	15	US-10-028-248A-40
2	98	27.8	1192	15	US-10-107-782-40
3	98	27.8	1394	9	US-09-842-930A-25
4	98	27.8	1416	14	US-10-133-172-4
5	98	27.8	1653	14	US-10-133-172-20
6	93	26.3	897	15	US-10-028-248A-41
7	93	26.3	897	15	US-10-107-782-41
8	93	26.3	897	15	US-10-107-782-211
9	93	26.3	1069	15	US-10-028-248A-39
10	93	26.3	1069	15	US-10-107-782-39
11	93	26.3	2420	15	US-10-028-248A-4
12	93	26.3	2420	15	US-10-107-782-4
13	93	26.3	2675	15	US-10-028-248A-2
14	93	26.3	2675	15	US-10-107-782-2
15	93	26.3	2675	15	US-10-107-782-2

Sequence 5, Appli
Sequence 2, Appli
Sequence 82, Appli
Sequence 17, Appli
Sequence 55, Appli
Sequence 43, Appli
Sequence 18, Appli

16 54 15.3 93 14 US-10-133-172-5
17 25 7.1 1431 9 US-09-842-930A-2
18 25 7.1 1431 14 US-10-133-172-2
19 20 5.7 106 14 US-10-023-896-82
20 18 5.1 18 14 US-10-133-172-17
21 15 4.2 15 9 US-09-842-930A-55
22 13 3.7 13 9 US-09-842-930A-43
23 13 3.7 13 14 US-10-133-172-18

ALIGNMENTS

RESULT 1
US-10-028-248A-40
; Sequence 40, Application US/10028248A
; Publication No. US20030235882A1

; GENERAL INFORMATION:

; APPLICANT: Shimkets, Richard
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corine
; APPLICANT: Casman, Stacie
; APPLICANT: Malyankar, Uriel
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Gangolli, Esha
; APPLICANT: Miller, Charles
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Smithson, Glennda
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Liu, Xiaohong
; APPLICANT: Colman, Steven
; APPLICANT: Tchernev, Velizar
; APPLICANT: Si, Jingsheng
; APPLICANT: Edinger, Shlomit
; APPLICANT: Stone, David
; APPLICANT: Sciore, Paul
; APPLICANT: Millet, Isabelle
; APPLICANT: Rothenberg, Mark
; TITLE OF INVENTION: No. US20030235882A1el Nucleic Acids and Polypeptides and Methods
; FILE REFERENCE: 21402-222
; CURRENT APPLICATION NUMBER: US/10/028,248A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256619
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262959
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311266
; NUMBER OF SEQ ID NOS: 211
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 40
; LENGTH: 1192
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-028-248A-40

Query Match 27.8%; Score 98; DB 15; Length 1192;
Best Local Similarity 100.0%; Pred. No. 1.8e-89;
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 SGNLLVMSFSLNFTLTVLAYNSARGAFLEHLDLSIRGTLFVPSNGLENET 228

Db 952 SGNLLQVLMSPFSLTNFLTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENET 1011
 QY 229 LSGRDIEHHLANSMFFYNDLVNGTTTLQTRLGSKLLIT 266
 Db 1012 LSGRDIEHHLANSMFFYNDLVNGTTTLQTRLGSKLLIT 1049

RESULT 2

US-10-107-782-40
 ; Sequence 40, Application US/10107782
 ; Publication No. US20040018970A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Boldog, Ferenc
 ; APPLICANT: Casman, Stacie
 ; APPLICANT: Coleman, Steve
 ; APPLICANT: Edinger, Shlomit
 ; APPLICANT: Gangolli, Esna
 ; APPLICANT: Kexuda, Ramesh
 ; APPLICANT: Li, Li
 ; APPLICANT: Liu, Xiaohong
 ; APPLICANT: Malyankar, Uriel
 ; APPLICANT: Miller, Charles
 ; APPLICANT: Millet, Isabelle
 ; APPLICANT: Patturajan, Meera
 ; APPLICANT: Rothenberg, Mark
 ; APPLICANT: Sciore, Paul
 ; APPLICANT: Shenoy, Suresh
 ; APPLICANT: Shimkets, Richard
 ; APPLICANT: Si, Jingsheng
 ; APPLICANT: Smithson, Glenda
 ; APPLICANT: Spytek, Kimberly
 ; APPLICANT: Stone, David
 ; APPLICANT: Taupier, Raymond, Jr.
 ; APPLICANT: Tchernev, Velizar
 ; APPLICANT: Vernet, Corine
 ; APPLICANT: Zerhusen, Brian
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES AND METHODS OF USE THEREOF
 ; FILE REFERENCE: 21402-222CIP
 ; CURRENT APPLICATION NUMBER: US/10/107,782
 ; CURRENT FILING DATE: 2002-03-27
 ; PRIOR FILING DATE: 2001-12-19
 ; PRIOR APPLICATION NUMBER: 10/028,248
 ; PRIOR FILING DATE: 2001-12-19
 ; PRIOR APPLICATION NUMBER: 60/256,619
 ; PRIOR FILING DATE: 2000-12-19
 ; PRIOR APPLICATION NUMBER: 60/262,959
 ; PRIOR FILING DATE: 2001-01-19
 ; PRIOR APPLICATION NUMBER: 60/272,408
 ; PRIOR FILING DATE: 2001-02-28
 ; PRIOR APPLICATION NUMBER: 60/285,189
 ; PRIOR FILING DATE: 2001-04-20
 ; PRIOR APPLICATION NUMBER: 60/308,039
 ; PRIOR FILING DATE: 2001-07-26
 ; PRIOR APPLICATION NUMBER: 60/311,266
 ; PRIOR FILING DATE: 2001-08-09
 ; PRIOR APPLICATION NUMBER: 60/279,344
 ; PRIOR FILING DATE: 2001-03-28
 ; NUMBER OF SEQ ID NOS: 215
 ; SOFTWARE: Curaseq1st version 0.1
 ; SEQ ID NO 40
 ; LENGTH: 1192
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-107-782-40

Query Match 27.8%; Score 98; DB 15; Length 1192;
 Best Local Similarity 100.0%; Pred. No. 1.8e-89;
 Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 169 SGNLLQVLMSPFSLTNFLTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENET 228
 Db 952 SGNLLQVLMSPFSLTNFLTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENET 1011
 QY 229 LSGRDIEHHLANSMFFYNDLVNGTTTLQTRLGSKLLIT 266

Db 1012 LSGRDIEHHLANSMFFYNDLVNGTTTLQTRLGSKLLIT 1049

RESULT 3

US-09-842-930A-25
 ; Sequence 25, Application US/09842930A
 ; Publication No. US20020197681A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Weigel, Paul
 ; TITLE OF INVENTION: Identification of Hyaluronan Receptor for Endocytosis
 ; FILE REFERENCE: 5820.603
 ; CURRENT APPLICATION NUMBER: US/09/842,930A
 ; CURRENT FILING DATE: 2001-04-22
 ; PRIOR APPLICATION NUMBER: 60/245,320
 ; PRIOR FILING DATE: 2000-11-02
 ; PRIOR APPLICATION NUMBER: 60/199,538
 ; PRIOR FILING DATE: 2000-04-25
 ; NUMBER OF SEQ ID NOS: 56
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 25
 ; LENGTH: 1394
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-842-930A-25

Query Match 27.8%; Score 98; DB 9; Length 1394;
 Best Local Similarity 100.0%; Pred. No. 2.1e-89;
 Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 169 SGNLLQVLMSPFSLTNFLTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENET 228
 Db 1154 SGNLLQVLMSPFSLTNFLTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENET 1213
 QY 229 LSGRDIEHHLANSMFFYNDLVNGTTTLQTRLGSKLLIT 266
 Db 1214 LSGRDIEHHLANSMFFYNDLVNGTTTLQTRLGSKLLIT 1251

RESULT 4

US-10-133-172-4
 ; Sequence 4, Application US/10133172
 ; Publication No. US20030104987A1
 ; GENERAL INFORMATION:
 ; APPLICANT: WEIGEL, PAUL H
 ; APPLICANT: WEIGEL, JANET A
 ; TITLE OF INVENTION: METHODS OF USING THE HYALURONAN RECEPTOR FOR ENDOCYTOSIS
 ; FILE REFERENCE: 5864.014
 ; CURRENT APPLICATION NUMBER: US/10/133,172
 ; CURRENT FILING DATE: 2002-04-25
 ; PRIOR APPLICATION NUMBER: 60/286,468
 ; PRIOR FILING DATE: 2001-04-25
 ; PRIOR APPLICATION NUMBER: 09/842,930
 ; PRIOR FILING DATE: 2001-04-25
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 4
 ; LENGTH: 1416
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-133-172-4

Query Match 27.8%; Score 98; DB 14; Length 1416;
 Best Local Similarity 100.0%; Pred. No. 2.1e-89;
 Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 169 SGNLLQVLMSPFSLTNFLTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENET 228
 Db 1176 SGNLLQVLMSPFSLTNFLTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENET 1235
 QY 229 LSGRDIEHHLANSMFFYNDLVNGTTTLQTRLGSKLLIT 266
 Db 1236 LSGRDIEHHLANSMFFYNDLVNGTTTLQTRLGSKLLIT 1273

RESULT 5
US-10-133-172-20
; Sequence 20, Application US/10133172
; Publication No. US20030104987A1
; GENERAL INFORMATION:
; APPLICANT: WEIGEL, PAUL H A
; APPLICANT: WEIGEL, JANET A
; TITLE OF INVENTION: METHODS OF USING THE HYALURONAN RECEPTOR FOR ENDOCYTOSIS
; FILE REFERENCE: 5864.014
; CURRENT APPLICATION NUMBER: US/10/133,172
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 60/286,468
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 09/842,930
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 1653
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-133-172-20

Query Match 27.8%; Score 98; DB 14; Length 1653;
Best Local Similarity 100.0%; Pred. No. 2.4e-89;
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 SGNLQVLMSPPSLTNPLTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENET 228
Db 1413 SGNLQVLMSPPSLTNPLTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENET 1472

QY 229 LSGRDIHHLANVMFFYNDLVNCTTLQTRLGSKLLIT 266
Db 1473 LSGRDIHHLANVMFFYNDLVNCTTLQTRLGSKLLIT 1510

RESULT 6
US-10-028-248A-41
; Sequence 41, Application US/10028248A
; Publication No. US20030235882A1
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corine
; APPLICANT: Casman, Stacie
; APPLICANT: Malyankar, Uriel
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Gangolli, Esha
; APPLICANT: Miller, Charles
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Smithson, Glennda
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Liu, Xiaohong
; APPLICANT: Colman, Steven
; APPLICANT: Tchernev, Velizar
; APPLICANT: Si, Jingsheng
; APPLICANT: Edinger, Shlomit
; APPLICANT: Stone, David
; APPLICANT: Sciore, Paul
; APPLICANT: Millet, Isabelle
; APPLICANT: Rothenberg, Mark
; TITLE OF INVENTION: No. US20030235882A1el Nucleic Acids and Polypeptides and Methods
; FILE REFERENCE: 21402-222
; CURRENT APPLICATION NUMBER: US/10/028,248A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256619
; PRIOR FILING DATE: 2001-12-19

; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262959
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311266
; PRIOR FILING DATE: 2001-08-09
; NUMBER OF SEQ ID NOS: 211
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 897
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-028-248A-41

Query Match 26.3%; Score 93; DB 15; Length 897;
Best Local Similarity 100.0%; Pred. No. 1.5e-84;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTGPGKHCKECKSHYVGDGLNCEPEQLPIDRCLQDNQGCHADAKCVDLHFQDTTGVGFHL 60
Db 489 MTGPGKHCKECKSHYVGDGLNCEPEQLPIDRCLQDNQGCHADAKCVDLHFQDTTGVGFHL 548

QY 61 RSPIGQVKLTDFDKAREACANEATMATYNOLSY 93
Db 549 RSPIGQVKLTDFDKAREACANEATMATYNOLSY 581

RESULT 7
US-10-028-248A-211
; Sequence 211, Application US/10028248A
; Publication No. US20030235882A1
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corine
; APPLICANT: Casman, Stacie
; APPLICANT: Malyankar, Uriel
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Gangolli, Esha
; APPLICANT: Miller, Charles
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Smithson, Glennda
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Liu, Xiaohong
; APPLICANT: Colman, Steven
; APPLICANT: Tchernev, Velizar
; APPLICANT: Si, Jingsheng
; APPLICANT: Edinger, Shlomit
; APPLICANT: Stone, David
; APPLICANT: Sciore, Paul
; APPLICANT: Millet, Isabelle
; APPLICANT: Rothenberg, Mark
; TITLE OF INVENTION: No. US20030235882A1el Nucleic Acids and Polypeptides and Methods
; FILE REFERENCE: 21402-222
; CURRENT APPLICATION NUMBER: US/10/028,248A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256619
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262959
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285189

; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311266
; PRIOR FILING DATE: 2001-08-09
; NUMBER OF SEQ ID NOS: 211
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 211
; LENGTH: 897
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-028-248A-211

Query Match 26.3%; Score 93; DB 15; Length 897;
Best Local Similarity 100.0%; Pred. No. 1.5e-84;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTGPGKHCKCKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVFHL 60
Db 489 MTGPGKHCKCKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVFHL 548
QY 61 RSPGQYKLTDFDKAREACANEATMATYNQLSY 93
Db 549 RSPGQYKLTDFDKAREACANEATMATYNQLSY 581

RESULT 8
US-10-107-782-41
; Sequence 41, Application US/10107782
; Publication No. US20040018970A1
; GENERAL INFORMATION:
; APPLICANT: Boldog, Ferenc,
; APPLICANT: Casman, Stacie,
; APPLICANT: Colman, Steve,
; APPLICANT: Edinger, Shlomit,
; APPLICANT: Gangolli, Esha,
; APPLICANT: Kekuda, Ramesh,
; APPLICANT: Li, Li,
; APPLICANT: Liu, Xiaohong,
; APPLICANT: Malyankar, Uriel,
; APPLICANT: Miller, Charles,
; APPLICANT: Millet, Isabelle,
; APPLICANT: Patturajan, Meera,
; APPLICANT: Rothenberg, Mark,
; APPLICANT: Sciore, Paul,
; APPLICANT: Shenoy, Suresh,
; APPLICANT: Shimkets, Richard,
; APPLICANT: Si, Jingsheng,
; APPLICANT: Smithson, Glennda,
; APPLICANT: Spytek, Kimberly,
; APPLICANT: Stone, David,
; APPLICANT: Taupier, Raymond, Jr.,
; APPLICANT: Tchernev, Velizar,
; APPLICANT: Vernet, Corine,
; APPLICANT: Zerhusen, Brian
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: 21402-222CIP
; CURRENT APPLICATION NUMBER: US/10/107,782
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: 10/028,248
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,619
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262,959
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272,408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285,189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308,039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311,266
; PRIOR FILING DATE: 2001-08-09

; PRIOR APPLICATION NUMBER: 60/279,344
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 215
; SOFTWARE: Curaseqlist version 0.1
; SEQ ID NO 41
; LENGTH: 897
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-107-782-41

Query Match 26.3%; Score 93; DB 15; Length 897;
Best Local Similarity 100.0%; Pred. No. 1.5e-84;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTGPGKHCKCKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVFHL 60
Db 489 MTGPGKHCKCKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVFHL 548
QY 61 RSPGQYKLTDFDKAREACANEATMATYNQLSY 93
Db 549 RSPGQYKLTDFDKAREACANEATMATYNQLSY 581

RESULT 9
US-10-107-782-211
; Sequence 211, Application US/10107782
; Publication No. US20040018970A1
; GENERAL INFORMATION:
; APPLICANT: Boldog, Ferenc,
; APPLICANT: Casman, Stacie,
; APPLICANT: Colman, Steve,
; APPLICANT: Edinger, Shlomit,
; APPLICANT: Gangolli, Esha,
; APPLICANT: Kekuda, Ramesh,
; APPLICANT: Li, Li,
; APPLICANT: Liu, Xiaohong,
; APPLICANT: Malyankar, Uriel,
; APPLICANT: Miller, Charles,
; APPLICANT: Millet, Isabelle,
; APPLICANT: Patturajan, Meera,
; APPLICANT: Rothenberg, Mark,
; APPLICANT: Sciore, Paul,
; APPLICANT: Shenoy, Suresh,
; APPLICANT: Shimkets, Richard,
; APPLICANT: Si, Jingsheng,
; APPLICANT: Smithson, Glennda,
; APPLICANT: Spytek, Kimberly,
; APPLICANT: Stone, David,
; APPLICANT: Taupier, Raymond, Jr.,
; APPLICANT: Tchernev, Velizar,
; APPLICANT: Vernet, Corine,
; APPLICANT: Zerhusen, Brian
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: 21402-222CIP
; CURRENT APPLICATION NUMBER: US/10/107,782
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: 10/028,248
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,619
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262,959
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272,408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285,189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308,039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311,266
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/279,344
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 215


```

; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 211
; LENGTH: 897
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-107-782-211

Query Match      26.3%; Score 93; DB 15; Length 897;
Best Local Similarity 100.0%; Pred. No. 1.5e-84;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTGPGKHCKECSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDITTVGVFHL 60
Db 489 MTGPGKHCKECSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDITTVGVFHL 548

Qy 61 RSPGQYKLTDFDKAREACANEAAATMATYNQLSY 93
Db 549 RSPGQYKLTDFDKAREACANEAAATMATYNQLSY 581

RESULT 10
US-10-028-248A-39
; Sequence 39, Application US/10028248A
; Publication No. US20030235882A1
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corine
; APPLICANT: Casman, Stacie
; APPLICANT: Malyankar, Uziel
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Gangolli, Esha
; APPLICANT: Miller, Charles
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Smithson, Glenda
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Liu, Xiaohong
; APPLICANT: Colman, Steven
; APPLICANT: Tchernev, Velizar
; APPLICANT: Si, Jingsheng
; APPLICANT: Edinger, Shlomit
; APPLICANT: Stone, David
; APPLICANT: Sciore, Paul
; APPLICANT: Millet, Isabelle
; APPLICANT: Rothenberg, Mark
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides and Methods
; FILE REFERENCE: 21402-222
; CURRENT FILING DATE: 2001-12-19
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256619
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262959
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311266
; NUMBER OF SEQ ID NOS: 211
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 39
; LENGTH: 1069
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-028-248A-39

```

```

Query Match      26.3%; Score 93; DB 15; Length 1069;
Best Local Similarity 100.0%; Pred. No. 1.8e-84;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTGPGKHCKECSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDITTVGVFHL 60
Db 697 MTGPGKHCKECSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDITTVGVFHL 756

Qy 61 RSPGQYKLTDFDKAREACANEAAATMATYNQLSY 93
Db 757 RSPGQYKLTDFDKAREACANEAAATMATYNQLSY 789

RESULT 11
US-10-107-782-39
; Sequence 39, Application US/10107782
; Publication No. US20040018970A1
; GENERAL INFORMATION:
; APPLICANT: Boldog, Ferenc
; APPLICANT: Casman, Stacie
; APPLICANT: Coleman, Steve
; APPLICANT: Edinger, Shlomit
; APPLICANT: Gangolli, Esha
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Li, Li
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uziel
; APPLICANT: Miller, Charles
; APPLICANT: Millet, Isabelle
; APPLICANT: Patturajan, Meera
; APPLICANT: Rothenberg, Mark
; APPLICANT: Sciore, Paul
; APPLICANT: Shenoy, Suresh
; APPLICANT: Shimkets, Richard
; APPLICANT: Si, Jingsheng
; APPLICANT: Smithson, Glenda
; APPLICANT: Spytek, Kimberly
; APPLICANT: Stone, David
; APPLICANT: Taupier, Raymond, Jr.
; APPLICANT: Tchernev, Velizar
; APPLICANT: Vernet, Corine
; APPLICANT: Zerhusen, Brian
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: 21402-222CIP
; CURRENT FILING DATE: 2002-03-27
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,619
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262,959
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272,408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285,189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308,039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311,266
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/279,344
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 215
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 39
; LENGTH: 1069
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-107-782-39

Query Match      26.3%; Score 93; DB 15; Length 1069;
Best Local Similarity 100.0%; Pred. No. 1.8e-84;

```

Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTGPGKHCKSHYVGDGNCPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVFHL 60

Db 697 MTGPGKHCKSHYVGDGNCPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVFHL 756

QY 61 RSLPGQYKLTDFDKAREACANEATMATYNQLSY 93

Db 757 RSLPGQYKLTDFDKAREACANEATMATYNQLSY 789

RESULT 12

US-10-028-248A-4

; Sequence 4, Application US/10028248A

; Publication No. US20030235882A1

; GENERAL INFORMATION:

; APPLICANT: Shinkets, Richard

; APPLICANT: Vernet, Corine

; APPLICANT: Patturajan, Meera

; APPLICANT: Casman, Stacie

; APPLICANT: Malyankar, Uriel

; APPLICANT: Shenoy, Suresh

; APPLICANT: Gangolli, Esha

; APPLICANT: Miller, Charles

; APPLICANT: Boldog, Ferenc

; APPLICANT: Li, Li

; APPLICANT: Taupier Jr, Raymond J

; APPLICANT: Kekuda, Ramesh

; APPLICANT: Smithson, Glennda

; APPLICANT: Zernhusen, Bryan

; APPLICANT: Liu, Xiaohong

; APPLICANT: Coleman, Steven

; APPLICANT: Tchernev, Velizar

; APPLICANT: Si, Jingsheng

; APPLICANT: Edinger Shlomit

; APPLICANT: Stone, David

; APPLICANT: Sciore, Paul

; APPLICANT: Miller, Isabelle

; APPLICANT: Rothenberg, Mark

; TITLE OF INVENTION: No. US20030235882A1el Nucleic Acids and Polypeptides and Methods

; FILE REFERENCE: 21402-222

; CURRENT APPLICATION NUMBER: US/10/028,248A

; CURRENT FILING DATE: 2001-12-19

; PRIOR APPLICATION NUMBER: 60/256619

; PRIOR FILING DATE: 2000-12-19

; PRIOR APPLICATION NUMBER: 60/262959

; PRIOR FILING DATE: 2001-01-19

; PRIOR APPLICATION NUMBER: 60/272408

; PRIOR FILING DATE: 2001-02-28

; PRIOR APPLICATION NUMBER: 60/285189

; PRIOR FILING DATE: 2001-04-20

; PRIOR APPLICATION NUMBER: 60/308039

; PRIOR FILING DATE: 2001-07-26

; PRIOR APPLICATION NUMBER: 60/311266

; PRIOR FILING DATE: 2001-08-09

; NUMBER OF SEQ ID NOS: 211

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 4

; LENGTH: 2420

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: UNSURE

; LOCATION: (371)..(372)

; OTHER INFORMATION: Wherein Xaa is any naturally occurring amino acid

US-10-028-248A-4

Query Match 26.3%; Score 93; DB 15; Length 2420;

Best Local Similarity 100.0%; Pred. No. 3.8e-84;

Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTGPGKHCKSHYVGDGNCPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVFHL 60

Db 2250 MTGPGKHCKSHYVGDGNCPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVFHL 2309

QY 61 RSLPGQYKLTDFDKAREACANEATMATYNQLSY 93

Db 2310 RSLPGQYKLTDFDKAREACANEATMATYNQLSY 2342

RESULT 13

US-10-107-782-4

; Sequence 4, Application US/10107782

; Publication No. US20040018970A1

; GENERAL INFORMATION:

; APPLICANT: Boldog, Ferenc

; APPLICANT: Casman, Stacie

; APPLICANT: Coleman, Steve

; APPLICANT: Edinger, Shlomit

; APPLICANT: Gangolli, Esha

; APPLICANT: Kekuda, Ramesh

; APPLICANT: Li, Li

; APPLICANT: Liu, Xiaohong

; APPLICANT: Malyankar, Uriel

; APPLICANT: Miller, Charles

; APPLICANT: Millet, Isabelle

; APPLICANT: Patturajan, Meera

; APPLICANT: Rothenberg, Mark

; APPLICANT: Sciore, Paul

; APPLICANT: Shenoy, Suresh

; APPLICANT: Shinkets, Richard

; APPLICANT: Si, Jingsheng

; APPLICANT: Smithson, Glennda

; APPLICANT: Spytek, Kimberly

; APPLICANT: Stone, David

; APPLICANT: Taupier, Raymond, jr.,

; APPLICANT: Tchernev, Velizar

; APPLICANT: Vernet, Corine

; APPLICANT: Zernhusen, Brian

; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES AND METHODS OF USE THEREOF

; FILE REFERENCE: 21402-222CIP

; CURRENT APPLICATION NUMBER: US/10/107,782

; CURRENT FILING DATE: 2002-03-27

; PRIOR APPLICATION NUMBER: 10/028,248

; PRIOR FILING DATE: 2001-12-19

; PRIOR APPLICATION NUMBER: 60/256,619

; PRIOR FILING DATE: 2000-12-19

; PRIOR APPLICATION NUMBER: 60/262,959

; PRIOR FILING DATE: 2001-01-19

; PRIOR APPLICATION NUMBER: 60/272,408

; PRIOR FILING DATE: 2001-02-28

; PRIOR APPLICATION NUMBER: 60/285,189

; PRIOR FILING DATE: 2001-04-20

; PRIOR APPLICATION NUMBER: 60/308,039

; PRIOR FILING DATE: 2001-07-26

; PRIOR APPLICATION NUMBER: 60/311,266

; PRIOR FILING DATE: 2001-08-09

; PRIOR APPLICATION NUMBER: 60/279,344

; PRIOR FILING DATE: 2001-03-28

; NUMBER OF SEQ ID NOS: 215

; SOFTWARE: CuraseqList version 0.1

; SEQ ID NO 4

; LENGTH: 2420

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (371)..(371)

; OTHER INFORMATION: Wherein Xaa may be any naturally occurring amino acid

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (372)..(372)

; OTHER INFORMATION: Wherein Xaa may be any naturally occurring amino acid

US-10-107-782-4

```

Query Match      26.3%; Score 93; DB 15; Length 2420;
Best Local Similarity 100.0%; Pred. No. 3.8e-84;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTGPGKHCKSHYVGDGLNCEPEQLPIDRCLQDNQCHADAKCVDLHFQDITTVGVFHL 60
    |||||
Db 2250 MTGPGKHCKSHYVGDGLNCEPEQLPIDRCLQDNQCHADAKCVDLHFQDITTVGVFHL 2309

QY 61 RSPLGQYKLTDFDKAREACANEATMATYNQLSY 93
    |||||
Db 2310 RSPLGQYKLTDFDKAREACANEATMATYNQLSY 2342

```

```

RESULT 14
US-10-028-248A-2
; Sequence 2, Application US/10028248A
; Publication No. US20030235882A1
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corine
; APPLICANT: Casman, Stacie
; APPLICANT: Malyankar, Uriel
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Gangolli, Esha
; APPLICANT: Miller, Charles
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Smithson, Glennnda
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Liu, Xiaohong
; APPLICANT: Colman, Steven
; APPLICANT: Tchernev, Velizar
; APPLICANT: Si, Jingsheng
; APPLICANT: Edinger, Shlomit
; APPLICANT: Stone, David
; APPLICANT: Sciore, Paul
; APPLICANT: Millet, Isabelle
; APPLICANT: Rothenberg, Mark
; TITLE OF INVENTION: No. US20030235882A1el Nucleic Acids and Polypeptides and Methods
; FILE REFERENCE: 21402-222
; CURRENT APPLICATION NUMBER: US/10/028,248A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256619
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262959
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311266
; PRIOR FILING DATE: 2001-08-09
; NUMBER OF SEQ ID NOS: 211
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2675
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (55)..(56)
; OTHER INFORMATION: Wherein Xaa is any naturally occurring amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (427)..(428)

```

```

; OTHER INFORMATION: Wherein Xaa is any naturally occurring amino acid
US-10-028-248A-2

Query Match      26.3%; Score 93; DB 15; Length 2675;
Best Local Similarity 100.0%; Pred. No. 4.1e-84;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTGPGKHCKSHYVGDGLNCEPEQLPIDRCLQDNQCHADAKCVDLHFQDITTVGVFHL 60
    |||||
Db 2303 MTGPGKHCKSHYVGDGLNCEPEQLPIDRCLQDNQCHADAKCVDLHFQDITTVGVFHL 2362

QY 61 RSPLGQYKLTDFDKAREACANEATMATYNQLSY 93
    |||||
Db 2363 RSPLGQYKLTDFDKAREACANEATMATYNQLSY 2395

RESULT 15
US-10-107-782-2
; Sequence 2, Application US/10107782
; Publication No. US20040018970A1
; GENERAL INFORMATION:
; APPLICANT: Boldog, Ferenc
; APPLICANT: Casman, Stacie
; APPLICANT: Colman, Steve
; APPLICANT: Edinger, Shlomit
; APPLICANT: Gangolli, Esha
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Li, Li
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel
; APPLICANT: Miller, Charles
; APPLICANT: Millet, Isabelle
; APPLICANT: Patturajan, Meera
; APPLICANT: Rothenberg, Mark
; APPLICANT: Sciore, Paul
; APPLICANT: Shenoy, Suresh
; APPLICANT: Shimkets, Richard
; APPLICANT: Si, Jingsheng
; APPLICANT: Smithson, Glennnda
; APPLICANT: Spytek, Kimberly
; APPLICANT: Stone, David
; APPLICANT: Taupier, Raymond, Jr.,
; APPLICANT: Tchernev, Velizar
; APPLICANT: Vernet, Corine
; APPLICANT: Zerhusen, Brian
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: 21402-222CIP
; CURRENT APPLICATION NUMBER: US/10/107,782
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: 10/028,248
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,619
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262,959
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272,408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285,189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308,039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311,266
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/279,344
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 215
; SOFTWARE: Curaseqlist version 0.1
; SEQ ID NO 2
; LENGTH: 2675
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature

```

```
; LOCATION: (55)..(55)
; OTHER INFORMATION: Wherein Xaa may be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (56)..(56)
; OTHER INFORMATION: Wherein Xaa may be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (427)..(427)
; OTHER INFORMATION: Wherein Xaa may be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (428)..(428)
; OTHER INFORMATION: Wherein Xaa may be any naturally occurring amino acid
US-10-107-782-2

Query Match          26.3%; Score 93; DB 15; Length 2675;
Best Local Similarity 100.0%; Pred. No. 4.1e-84;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTGPGKHCKECSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTVGVFHL 60
Db 2303 MTGPGKHCKECSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTVGVFHL 2362

Qy 61 RSLPGQKLTDPDKAREACANEATMATVNLQSY 93
Db 2363 RSLPGQKLTDPDKAREACANEATMATVNLQSY 2395

RESULT 16
US-10-133-172-5
; Sequence 5, Application US/10133172
; Publication No. US20030104987A1
; GENERAL INFORMATION:
; APPLICANT: WEIGEL, PAUL H
; TITLE OF INVENTION: METHODS OF USING THE HYALURONAN RECEPTOR FOR ENDOCYTOSIS
; FILE REFERENCE: 5864.014
; CURRENT APPLICATION NUMBER: US/10/133,172
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 60/286,468
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 09/842,930
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-133-172-5

Query Match          15.3%; Score 54; DB 14; Length 93;
Best Local Similarity 100.0%; Pred. No. 3.8e-46;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 95 QKAKYHLCAGWLETGRVAYPTAFASQCGSGVGVIVDYGPRPNKSEMDVFCY 148
Db 40 QKAKYHLCAGWLETGRVAYPTAFASQCGSGVGVIVDYGPRPNKSEMDVFCY 93

RESULT 17
US-09-842-930A-2
; Sequence 2, Application US/09842930A
; Publication No. US20020197681A1
; GENERAL INFORMATION:
; APPLICANT: Weigel, Paul
; TITLE OF INVENTION: Identification of Hyaluronan Receptor for Endocytosis
; FILE REFERENCE: 5820.503
; CURRENT APPLICATION NUMBER: US/09/842,930A
; CURRENT FILING DATE: 2001-04-22
; PRIOR APPLICATION NUMBER: 60/245,320
; PRIOR FILING DATE: 2000-11-02
```

```
; PRIOR APPLICATION NUMBER: 60/199,538
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1431
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-842-930A-2

Query Match          7.1%; Score 25; DB 9; Length 1431;
Best Local Similarity 100.0%; Pred. No. 7.6e-16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 FQDTTVGVFHLRSLPGQKLTDFDKA 74
Db 1057 FQDTTVGVFHLRSLPGQKLTDFDKA 1081

RESULT 18
US-10-133-172-2
; Sequence 2, Application US/10133172
; Publication No. US20030104987A1
; GENERAL INFORMATION:
; APPLICANT: WEIGEL, JANET A
; TITLE OF INVENTION: METHODS OF USING THE HYALURONAN RECEPTOR FOR ENDOCYTOSIS
; FILE REFERENCE: 5864.014
; CURRENT APPLICATION NUMBER: US/10/133,172
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 60/286,468
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 09/842,930
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1431
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-133-172-2

Query Match          7.1%; Score 25; DB 14; Length 1431;
Best Local Similarity 100.0%; Pred. No. 7.6e-16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 FQDTTVGVFHLRSLPGQKLTDFDKA 74
Db 1057 FQDTTVGVFHLRSLPGQKLTDFDKA 1081

RESULT 19
US-10-023-896-82
; Sequence 82, Application US/10023896
; Publication No. US20030027776A1
; GENERAL INFORMATION:
; APPLICANT: Victor Roschke
; TITLE OF INVENTION: 29 Human Cancer Associated Proteins
; FILE REFERENCE: PA004p1
; CURRENT APPLICATION NUMBER: US/10/023,896
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: unassigned
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: PCT/US00/23794
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/152,296
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: 60/158,003
; PRIOR FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 82
; LENGTH: 106
```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-896-82

Query Match      5.7%; Score 20; DB 14; Length 106;
Best Local Similarity 100.0%; Pred. No. 7.8e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 326 IILVTGAVALAAYSYFRINR 345
Db 23 IILVTGAVALAAYSYFRINR 42

RESULT 20
US-10-133-172-17
; Sequence 17, Application US/10133172
; Publication No. US20030104987A1
; GENERAL INFORMATION:
; APPLICANT: WEIGEL, PAUL H
; TITLE OF INVENTION: IDENTIFICATION OF USING THE HYALURONAN RECEPTOR FOR ENDOCYTOSIS
; FILE REFERENCE: 5864.014
; CURRENT APPLICATION NUMBER: US/10/133,172
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 60/286,468
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 09/842,930
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-133-172-17

Query Match      5.1%; Score 18; DB 14; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.6e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TGPGRKCKCKSHVVGDG 19
Db 1 TGPGRKCKCKSHVVGDG 18

RESULT 21
US-09-842-930A-55
; Sequence 55, Application US/09842930A
; Publication No. US20020197681A1
; GENERAL INFORMATION:
; APPLICANT: Weigel, Paul
; TITLE OF INVENTION: Identification of Hyaluronan Receptor for Endocytosis
; FILE REFERENCE: 5820.603
; CURRENT APPLICATION NUMBER: US/09/842,930A
; CURRENT FILING DATE: 2001-04-22
; PRIOR APPLICATION NUMBER: 60/245,320
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: 60/199,538
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-842-930A-55

Query Match      4.2%; Score 15; DB 9; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 61 RSPLGQYKLTFFDKAR 75

```

```

Db 1 RSPLGQYKLTFFDKAR 15

RESULT 22
US-09-842-930A-43
; Sequence 43, Application US/09842930A
; Publication No. US20020197681A1
; GENERAL INFORMATION:
; APPLICANT: Weigel, Paul
; TITLE OF INVENTION: Identification of Hyaluronan Receptor for Endocytosis
; FILE REFERENCE: 5820.603
; CURRENT APPLICATION NUMBER: US/09/842,930A
; CURRENT FILING DATE: 2001-04-22
; PRIOR APPLICATION NUMBER: 60/245,320
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: 60/199,538
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-842-930A-43

Query Match      3.7%; Score 13; DB 9; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 213 GTLFVPQNSGLGE 225
Db 1 GTLFVPQNSGLGE 13

RESULT 23
US-10-133-172-18
; Sequence 18, Application US/10133172
; Publication No. US20030104987A1
; GENERAL INFORMATION:
; APPLICANT: WEIGEL, PAUL H
; APPLICANT: WEIGEL, JANET A
; TITLE OF INVENTION: METHODS OF USING THE HYALURONAN RECEPTOR FOR ENDOCYTOSIS
; FILE REFERENCE: 5864.014
; CURRENT APPLICATION NUMBER: US/10/133,172
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 60/286,468
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 09/842,930
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-133-172-18

Query Match      3.7%; Score 13; DB 14; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 28 PIDRCLQDNGQCH 40
Db 1 PIDRCLQDNGQCH 13

Search completed: April 1, 2004, 10:04:13
Job time : 46 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 09:55:43 ; Search time 22 Seconds
(without alignments)
828.362 Million cell updates/sec

Title: US-09-466-778B-11
Perfect score: 353
Sequence: 1 MTGPGKHCKCKSHYVDGL.....ALAAYSYFRINRKTIGFXHF 353

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 389414 seqs, 51625971 residues

Word size: 9

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 500 summaries

Database : Issued Patents AA.*
1: /cgm2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgm2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgm2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgm2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgm2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgm2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description

No matches found

Search completed: April 1, 2004, 09:59:31
Job time : 22 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 09:55:42 ; Search time 58 seconds
(without alignments)
1719.643 Million cell updates/sec

Title: US-09-466-778B-11

Perfect score: 353

Sequence: 1 MTGPGKHKCECKSHYVDGL.....ALAAYSYFRINRKTIGXFHF 353

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1586107 seqs, 282547505 residues

Word size: 9

Total number of hits satisfying chosen parameters: 20

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 500 summaries

Database: A Geneseq_29Jan04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	345	97.7	333	3	AAY93913
2	98	27.8	1394	5	AAM47684
3	98	27.8	1416	6	ABG72499
4	98	27.8	1653	6	ABG72514
5	93	26.3	244	4	AAB83359
6	93	26.3	669	4	AAB83364
7	93	26.3	689	4	AAB83362
8	93	26.3	897	5	ABJ10588
9	93	26.3	2420	5	ABJ10587
10	93	26.3	2675	5	ABJ10586
11	85	24.1	330	3	AAB42164
12	57	16.1	315	4	AAB83358
13	54	15.3	93	6	ABG72500
14	25	7.1	1431	5	AAM47675
15	25	7.1	1431	6	ABG72498
16	20	5.7	106	4	AUU00893
17	18	5.1	18	6	ABG72512
18	15	4.2	15	5	AAM47702
19	13	3.7	13	5	AAM47690
20	13	3.7	13	6	ABG72513

ALIGNMENTS

RESULT 1
AAY93913
ID AAY93913 standard; protein; 353 AA.
XX
AC AAY93913;
XX
DT 03-OCT-2000 (first entry)
XX
DE A human hyaluronan-binding protein, designated BM-HABP.
XX
KW Hyaluronan-binding protein WF-HABP; OE-HABP; BM-HABP;
KW proliferative condition; metastasis; inflammation; ischemia;
KW host defence dysfunction; immune surveillance dysfunction; arthritis;
KW multiple sclerosis; autoimmunity; immune dysfunction; allergy.
XX
OS Homo sapiens.
XX
FH Key
FT Misc-difference 94 /note= "unspecified amino acid encoded by NNC"
FT Domain 121..215 /note= "HA binding domain"
FT Misc-difference 157 /note= "unspecified amino acid encoded by TNC"
FT Misc-difference 303 /note= "unspecified amino acid encoded by YCT"
FT Misc-difference 314 /note= "unspecified amino acid encoded by NCC"
FT Misc-difference 320 /note= "unspecified amino acid encoded by GNA"
FT Misc-difference 324 /note= "unspecified amino acid encoded by TNT"
FT Misc-difference 325 /note= "unspecified amino acid encoded by GNC"
FT Misc-difference 351 /note= "unspecified amino acid encoded by CAN"
XX
PN WO200039166-A1.
XX
PD 06-JUL-2000.
XX
PF 20-DEC-1999; 99WO-US030462.
XX
PR 23-DEC-1998; 98US-0113871P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (AMNA-) AMERICAN NAT RED CROSS.
XX
FI Hastings GA, Liau G, Tsifrina E;
XX
WPI; 2000-452376/39.
DR N-PSDB; AAM57365.
XX
FT New hyaluron-binding proteins, known as full-length WF-HABP, WF-HABP, OE-HABP and BM-HABP, useful for treating proliferative conditions, metastasis, inflammation, ischemia, arthritis and multiple sclerosis.
XX
PS Claim 11; Fig 4A-B; 457pp; English.
XX
CC The present sequence represents a hyaluronan-binding protein. The specification describes four hyaluronan-binding proteins, known as WF-HABP, WF-HABP, OE-HABP, and BM-HABP. The polypeptides are useful for treating diseases such as proliferative conditions, metastasis, inflammation, ischemia, host defence dysfunction, immune surveillance dysfunction, arthritis, multiple sclerosis, autoimmunity, immune dysfunction and allergy
XX
SQ Sequence 353 AA;
Query Match 97.7%; Score 345; DB 3; Length 353;
Best Local Similarity 100.0%; Pred.No. 0;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTGPGKHCKESHYVGDGLNCPPEQLPIDRCLQDNGOCHADAKCVDLHFQDTTGVFHL 60
Db 1 MTGPGKHCKESHYVGDGLNCPPEQLPIDRCLQDNGOCHADAKCVDLHFQDTTGVFHL 60
QY 61 RSLPGQKLTDFDKAREACANEATMATYNOLSYXQAKYHLCAGWLETGRVAYPTAFAS 120
Db 61 RSLPGQKLTDFDKAREACANEATMATYNOLSYXQAKYHLCAGWLETGRVAYPTAFAS 120
QY 121 QNCGSGVVGIVDYGPRNPKSEMDVFCYRMKDVNCTKVGVDGSGFSYNLLOVLMSPF 180
Db 121 QNCGSGVVGIVDYGPRNPKSEMDVFCYRMKDVNCTKVGVDGSGFSYNLLOVLMSPF 180
QY 181 SLTNFLTEVLAYSNSSARGRAFLHLEHLDLSIRGTLFVPQNSGLGENETLSGRDIEHLAN 240
Db 181 SLTNFLTEVLAYSNSSARGRAFLHLEHLDLSIRGTLFVPQNSGLGENETLSGRDIEHLAN 240
QY 241 VSMFFYNDLVNGTTLQRLGSKLLITDRQDPLHPTETRCVDGRDITLWDICASNGITHVI 300
Db 241 VSMFFYNDLVNGTTLQRLGSKLLITDRQDPLHPTETRCVDGRDITLWDICASNGITHVI 300
QY 301 SRXIKAPPAVTLXHTGLGXGIFXXIILVTGAVALAAYSFRINRKTIGFXHF 353
Db 301 SRXIKAPPAVTLXHTGLGXGIFXXIILVTGAVALAAYSFRINRKTIGFXHF 353

RESULT 2
AAM47684
ID AAM47684 standard; protein; 1394 AA.
AC AAM47684;
XX
DT 22-FEB-2002 (first entry)
XX
DE Human Hyaluronic Acid Receptor for Endocytosis, HARE.
XX
KW HARE; Hyaluronic Acid Receptor for Endocytosis; hyaluronan; chondroitin;
KW chondroitin sulphate; extracellular matrix; cartilage; skin;
KW vitreous humour; endocytic receptor; glycosaminoglycan; human.
XX
OS Homo sapiens.
XX
PN WO200181544-A2.
XX
PD 01-NOV-2001.
XX
PF 25-APR-2001; 2001WO-US013403.
XX
PR 25-APR-2000; 2000US-0199538P.
PR 02-NOV-2000; 2000US-0245320P.
XX
PA (WEIG/) WEIGEL P A.
PA (ZHOU/) ZHOU B.
PA (WEIG/) WEIGEL J A.
XX
PI Weigel PA, Zhou B, Weigel JA;
XX
WPI; 2002-049271/06.
XX

New mammalian hyaluronic acid receptor for endocytosis, useful e.g. for identifying agents that inhibit binding to hyaluronic acid, and related nucleic acid.
XX
PS Claim 20; Fig 33; 263pp; English.
XX
CC The present invention relates to sequences for rat and human HARE
CC (Hyaluronic Acid Receptor for Endocytosis, ABA04548, ABA04562, AAM47675
CC and AAM47684). HARE can bind specifically to at least one of hyaluronic
CC acid (HA, also known as hyaluronan), chondroitin (C) or chondroitin
CC sulphate (CS). HA is an extracellular matrix component of all tissues, in
CC particular cartilage, skin and vitreous humour. HARE is the endocytic
CC receptor responsible for removing HA and other glycosaminoglycans from
CC the circulation. The present sequence is human HARE
XX

SQ Sequence 1394 AA;
Query Match 27.8%; Score 98; DB 5; Length 1394;
Best Local Similarity 100.0%; Pred. No. 3.4e-87;
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 169 SGNLQVLMSPFSLTNFLTEVLAYSNSSARGRAFLHLEHLDLSIRGTLFVPQNSGLGENET 228
Db 1154 SGNLQVLMSPFSLTNFLTEVLAYSNSSARGRAFLHLEHLDLSIRGTLFVPQNSGLGENET 1213
QY 229 LSGRDIIEHLANVSMFFYNDLVNGTTLQRLGSKLLIT 266
Db 1214 LSGRDIIEHLANVSMFFYNDLVNGTTLQRLGSKLLIT 1251

RESULT 3
ABG72499
ID ABG72499 standard; protein; 1416 AA.
XX
AC ABG72499;
XX
DT 19-FEB-2003 (first entry)
XX
DE Human 190kDa Hyaluronan receptor for endocytosis (HARE) #1.
XX
KW Hyaluronan receptor for endocytosis; HARE; hyaluronan; HA; CD;
KW chondroitin molecule; chondroitin sulphate; CDS; targeting; metastasis;
KW tumour; gene therapy; human; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200286093-A2.
XX
PD 31-OCT-2002.
XX
PF 25-APR-2002; 2002WO-US013209.
XX
PR 25-APR-2001; 2001US-00842930.
PR 25-APR-2001; 2001US-0286468P.
XX
PA (WEIG/) WEIGEL P H.
PA (WEIG/) WEIGEL J A.
XX
PI Weigel PH, Weigel JA;
XX
WPI; 2003-093126/08.
XX
N-PSDB; ABX13822.
XX

Targeting compounds e.g. chemotherapeutic agent to cell of subject expressing functional active hyaluronan receptor for endocytosis of HARE, by using hyaluronan, HARE, and/or monoclonal antibody that binds to HARE epitope.
XX
PS Example; Fig 9A; 167pp; English.
XX
CC The invention describes a method of targeting a compound to a cell or tissue of an individual expressing a functionally active hyaluronan (HA) receptor for endocytosis (HARE) or a cell that does not express functionally active HARE. The method involves using HA molecule, a chondroitin molecule (CD), a chondroitin sulphate (CDS) molecule, and/or a monoclonal antibody raised against a HA-binding domain of HARE. The method is useful for targeting a compound, preferably a chemotherapeutic agent or a radioisotope to cell of an individual, especially a human, expressing HARE on its surface (e.g. gene therapy). Also described is a method useful for preventing interaction between a cell having at least one of a HA coat, a CD coat and CDS coat and a cell expressing HARE on its surface. This second method is useful for preventing metastasis by preventing interaction between tumour cells having HA, CD or CDS coat and non-tumour cells expressing HARE on its surface. The invention also describes a method useful for detecting the presence of HA, CD and CDS in a biological fluid. This sequence encodes the human 190kDa Hyaluronan receptor for endocytosis (HARE)
XX

SQ Sequence 1416 AA;
 Query Match 27.8%; Score 98; DB 6; Length 1416;
 Best Local Similarity 100.0%; Pred. No. 3.4e-87;
 Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 169 SGNLLQVIMSPFSLTNFLTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENET 228
 Db 1176 SGNLLQVIMSPFSLTNFLTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENET 1235
 QY 229 LSGRDIEHHLANVSMFFYNDLVNGTTTLQTRLGSKLLIT 266
 Db 1236 LSGRDIEHHLANVSMFFYNDLVNGTTTLQTRLGSKLLIT 1273
 RESULT 4
 ABG72514
 ID ABG72514 standard; protein; 1653 AA.
 XX AC ABG72514;
 XX DT 19-FEB-2003 (first entry)
 XX DE Human 190kDa Hyaluronan receptor for endocytosis (HARE) #2.
 XX KW Hyaluronan receptor for endocytosis; HARE; hyaluronan; HA; CD;
 KW chondroitin molecule; chondroitin sulphate; CDS; targeting; metastasis;
 KW tumour; gene therapy; human.
 XX OS Homo sapiens.
 XX PN WO200286093-A2.
 XX PD 31-OCT-2002.
 XX PF 25-APR-2002; 2002WO-US013209.
 XX PR 25-APR-2001; 2001US-00842930.
 XX PR 25-APR-2001; 2001US-0286468P.
 XX PA (WEIG/) WEIGEL P H.
 XX PA (WEIG/) WEIGEL J A.
 XX PI Weigel PH, Weigel JA;
 XX DR WPI; 2003-093126/08.
 XX DR N-PSDB; ABX13823.
 XX PT Targeting compounds e.g. chemotherapeutic agent to cell of subject
 PT expressing functional active hyaluronan receptor for endocytosis of HARE,
 PT by using hyaluronan, HARE, and/or monoclonal antibody that binds to HARE
 PT epitope.
 XX PS Example; Fig 9B; 167pp; English.
 XX CC The invention describes a method of targeting a compound to a cell or
 CC tissue of an individual expressing a functionally active hyaluronan (HA)
 CC receptor for endocytosis (HARE) or a cell that does not express
 CC functionally active HARE. The method involves using HA molecule, a
 CC chondroitin molecule (CD), a chondroitin sulphate (CDS) molecule, and/or
 CC a monoclonal antibody raised against a HA-binding domain of HARE. The
 CC method is useful for targeting a compound, preferably a chemotherapeutic
 CC agent or a radioisotope to cell of an individual, especially a human,
 CC expressing HARE on its surface (e.g. gene therapy). Also described is a
 CC method useful for preventing interaction between a cell having at least
 CC one of a HA coat, a CD coat and CDS coat and a cell expressing HARE on
 CC its surface. This second method is useful for preventing metastasis by
 CC preventing interaction between tumour cells having HA, CD or CDS coat and
 CC non-tumour cells expressing HARE on its surface. The invention also
 CC describes a method useful for detecting the presence of HA, CD and CDS in
 CC a biological fluid. This is the amino acid sequence of a longer version
 CC of the human 190kDa Hyaluronan receptor for endocytosis (HARE) shown in
 CC ABG72499 (Encoded by ABX13822)

XX SQ Sequence 1653 AA;
 Query Match 27.8%; Score 98; DB 6; Length 1653;
 Best Local Similarity 100.0%; Pred. No. 3.9e-87;
 Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 169 SGNLLQVIMSPFSLTNFLTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENET 228
 Db 1413 SGNLLQVIMSPFSLTNFLTEVLAYSNSARGRAFLHLDLSIRGTLFVPQNSGLGENET 1472
 QY 229 LSGRDIEHHLANVSMFFYNDLVNGTTTLQTRLGSKLLIT 266
 Db 1473 LSGRDIEHHLANVSMFFYNDLVNGTTTLQTRLGSKLLIT 1510
 RESULT 5
 AAB83359
 ID AAB83359 standard; protein; 244 AA.
 XX AC AAB83359;
 XX DT 26-MAR-2002 (first entry)
 XX DE NOV3 protein sequence.
 XX KW NOV; Cytostatic; contraceptive; antiinflammatory; immunomodulatory;
 KW cardiovascular; casein kinase II phosphorylation site; contraception;
 KW serine/threonine kinase; Peutz-Jeghers syndrome; cellular proliferation;
 KW epidermal growth factor; cell development; apoptosis; cell adhesion;
 KW growth migration; cell structure; motility; cancer; immune disorder;
 KW inflammatory disorder; cellular adhesion disorder; long-QT syndrome;
 KW cardiovascular disease; hypertrophic cardiomyopathy; marfan syndrome;
 KW therapy; NOV1; NOV2; NOV3; NOV4; NOV5; NOV6; NOV7; NOV8; NOV9; NOV10;
 KW NOV11; NOV12; NOV13; NOV14; NOV15; NOV16.
 XX OS Unidentified.
 XX PN WO200136638-A2.
 XX PD 25-MAY-2001.
 XX PF 17-NOV-2000; 2000WO-US031543.
 XX PR 19-NOV-1999; 99US-0166336P.
 XX PR 29-NOV-1999; 99US-0167785P.
 XX PR 08-MAR-2000; 2000US-0187844P.
 XX PR 16-NOV-2000; 2000US-00715417.
 XX PA (CURA-) CURAGEN CORP.
 XX PI Shimkets RA, Lichenstein H, Vernet C, Fernandes E;
 XX DR WPI; 2001-648134/74.
 XX DR N-PSDB; AAF87114.
 XX PT Novel human polypeptides and the nucleic acids that encode them useful
 PT for preventing, diagnosing and treating e.g. cancer, inflammation and
 PT immune disorders.
 XX PS Claim 1; Page 14-17; 141pp; English.
 XX CC This sequence is the NOV3 protein. The invention relates to the NOV1-
 CC NOV16 proteins, and their coding sequences. The proteins have Cytostatic;
 CC contraceptive; antiinflammatory; immunomodulatory; and cardiovascular
 CC activities. The sequences may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate NOVX expression. They
 CC may be used to treat disorders associated with decreased expression by
 CC rectifying mutations or deletions in a patient's genome that affect the
 CC activity of protein by expressing inactive proteins or to supplement the
 CC patients own production of protein. They are used to produce NOVX
 CC proteins, by inserting the nucleic acid into a cell and culturing it to
 CC express the protein. The DNA may be used as DNA probes in assays to

CC detect and quantitate the presence of similar DNAs in samples, and which
 CC patients may need restorative therapy. The NOVX protein may also be used
 CC as antigens in the production of antibodies (Abs) against NOVX and in
 CC assays to identify modulators of NOVX expression and activity. The anti-
 CC NOVX Abs and antagonist are used to down regulate expression and
 CC activity. The anti-NOVX Abs are used for detecting the presence of NOVX
 CC in samples. Disorders that may be prevented, diagnosed and/or treated
 CC vary depending on the NOVX protein. NOV1, NOV3, NOV5, NOV7, NOV9-11 and
 CC NOV13-16 have casein kinase II phosphorylation sites characteristic of
 CC serine/threonine kinases, and are used to treat kinase-related disorders
 CC (e.g. Peutz-Jeghers syndrome, cellular proliferation and contraception).
 CC NOV2-3, NOV6 and NOV8 are homologous to the epidermal growth factor (EGF)
 CC -like super family and are involved in, e.g. regulation of cell
 CC development, apoptosis, cell adhesion, growth migration, cell structure
 CC and motility and protein management, and are used to treat cancers,
 CC inflammatory disorders, immune disorders and cellular adhesion disorders.
 CC NOV6-10 are homologous to EGF-like fibrillin proteins and are used to
 CC treat cardiovascular disease e.g. hypertrophic cardiomyopathy, long-QT
 CC syndrome and marfan syndrome
 CC
 CC Sequence 244 AA;

Query Match 26.3%; Score 93; DB 4; Length 244;
 Best Local Similarity 100.0%; Pred. No. 6.6e-83;
 Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTGPGKHCKCKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTVGVFHL 60
 Db 145 MTGPGKHCKCKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTVGVFHL 204
 QY 61 RSPGLQYKLTDFDKAREACANEATMATYNQLSY 93
 Db 205 RSPGLQYKLTDFDKAREACANEATMATYNQLSY 237

RESULT 6
 AAB83364
 ID AAB83364 standard; protein; 669 AA.
 AC AAB83364;
 XX
 XX

26-MAR-2002 (first entry)

NOV8 protein sequence.

NOV; Cytostatic; contraceptive; antiinflammatory; immunomodulatory;
 cardiovascular; casein kinase II phosphorylation site; contraception;
 serine/threonine kinase; Peutz-Jeghers syndrome; cellular proliferation;
 epidermal growth factor; cell development; apoptosis; cell adhesion;
 growth migration; cell structure; motility; cancer; immune disorder;
 inflammatory disorder; cellular adhesion disorder; long-QT syndrome;
 cardiovascular disease; hypertrophic cardiomyopathy; marfan syndrome;
 therapy; NOV1; NOV2; NOV3; NOV4; NOV5; NOV6; NOV7; NOV8; NOV9; NOV10;
 NOV11; NOV12; NOV13; NOV14; NOV15; NOV16.

Unidentified.

WO200136638-A2.

25-MAY-2001.

17-NOV-2000; 2000WO-US031543.

19-NOV-1999; 99US-0166336P.

29-NOV-1999; 99US-0167785P.

08-MAR-2000; 2000US-0187844P.

16-NOV-2000; 2000US-00715417.

(CURA-) CURAGEN CORP.

Shimkets RA, Lichenstein H, Vernet C, Fernandes E;
 WPI; 2001-648134/74.

DR N-PSDB; AAF971119.

XX Novel human polypeptides and the nucleic acids that encode them useful
 PT for preventing, diagnosing and treating e.g. cancer, inflammation and
 PT immune disorders.

XX Claim 1; Page 29-30; 141pp; English.

XX This sequence is the NOV8 protein. The invention relates to the NOV1-
 CC NOV16 proteins, and their coding sequences. The proteins have Cytostatic;
 CC contraceptive; antiinflammatory; immunomodulatory; and cardiovascular
 CC activities. The sequences may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate NOVX expression. They
 CC may be used to treat disorders associated with decreased expression by
 CC rectifying mutations or deletions in a patient's genome that affect the
 CC activity of protein by expressing inactive proteins or to supplement the
 CC patients own production of protein. They are used to produce NOVX
 CC proteins, by inserting the nucleic acid into a cell and culturing it to
 CC express the protein. The DNA may be used as DNA probes in assays to
 CC detect and quantitate the presence of similar DNAs in samples, and which
 CC patients may need restorative therapy. The NOVX protein may also be used
 CC as antigens in the production of antibodies (Abs) against NOVX and in
 CC assays to identify modulators of NOVX expression and activity. The anti-
 CC NOVX Abs and antagonist are used to down regulate expression and
 CC activity. The anti-NOVX Abs are used for detecting the presence of NOVX
 CC in samples. Disorders that may be prevented, diagnosed and/or treated
 CC vary depending on the NOVX protein. NOV1, NOV3, NOV5, NOV7, NOV9-11 and
 CC NOV13-16 have casein kinase II phosphorylation sites characteristic of
 CC serine/threonine kinases, and are used to treat kinase-related disorders
 CC (e.g. Peutz-Jeghers syndrome, cellular proliferation and contraception).
 CC NOV2-3, NOV6 and NOV8 are homologous to the epidermal growth factor (EGF)
 CC -like super family and are involved in, e.g. regulation of cell
 CC development, apoptosis, cell adhesion, growth migration, cell structure
 CC and motility and protein management, and are used to treat cancers,
 CC inflammatory disorders, immune disorders and cellular adhesion disorders.
 CC NOV6-10 are homologous to EGF-like fibrillin proteins and are used to
 CC treat cardiovascular disease e.g. hypertrophic cardiomyopathy, long-QT
 CC syndrome and marfan syndrome
 CC
 CC Sequence 669 AA;

Query Match 26.3%; Score 93; DB 4; Length 669;

Best Local Similarity 100.0%; Pred. No. 1.5e-82;

Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTGPGKHCKCKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTVGVFHL 60
 Db 566 MTGPGKHCKCKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTVGVFHL 625

QY 61 RSPGLQYKLTDFDKAREACANEATMATYNQLSY 93

Db 626 RSPGLQYKLTDFDKAREACANEATMATYNQLSY 658

RESULT 7

AAB83362

ID AAB83362 standard; protein; 669 AA.

XX AAB83362;

26-MAR-2002 (first entry)

NOV6 protein sequence.

NOV; Cytostatic; contraceptive; antiinflammatory; immunomodulatory;
 cardiovascular; casein kinase II phosphorylation site; contraception;
 serine/threonine kinase; Peutz-Jeghers syndrome; cellular proliferation;
 epidermal growth factor; cell development; apoptosis; cell adhesion;
 growth migration; cell structure; motility; cancer; immune disorder;
 inflammatory disorder; cellular adhesion disorder; long-QT syndrome;
 cardiovascular disease; hypertrophic cardiomyopathy; marfan syndrome;
 therapy; NOV1; NOV2; NOV3; NOV4; NOV5; NOV6; NOV7; NOV8; NOV9; NOV10;
 NOV11; NOV12; NOV13; NOV14; NOV15; NOV16.

XX OS Unidentified.
XX PN WO200136638-A2.
XX XX 25-MAY-2001.
XX XX 17-NOV-2000; 2000WO-US031543.
XX PR 19-NOV-1999; 98US-0166336P.
XX PR 29-NOV-1999; 99US-0167789P.
XX PR 08-MAR-2000; 2000US-0187844P.
XX PR 16-NOV-2000; 2000US-00715417.
XX XX (CURA-) CURAGEN CORP.
XX PI Shimkets RA, Lichenstein H, Vernet C, Fernandes E;
XX DR WPI; 2001-648134/74.
XX DR N-PSDB; AAF87117.
XX PT Novel human polypeptides and the nucleic acids that encode them useful
XX PT for preventing, diagnosing and treating e.g. cancer, inflammation and
XX PT immune disorders.
XX PS Claim 1; Page 24-25; 141pp; English.
XX CC This sequence is the NOV6 protein. The invention relates to the NOV1-
XX CC NOV6 proteins, and their coding sequences. The proteins have cytostatic;
XX CC contraceptive; antiinflammatory; immunomodulatory; and cardiovascular
XX CC activities. The sequences may be used in the prevention, diagnosis and
XX CC treatment of diseases associated with inappropriate NOVX expression. They
XX CC may be used to treat disorders associated with decreased expression by
XX CC rectifying mutations or deletions in a patient's genome that affect the
XX CC activity of protein by expressing inactive proteins or to supplement the
XX CC patients own production of protein. They are used to produce NOVX
XX CC proteins, by inserting the nucleic acid into a cell and culturing it to
XX CC express the protein. The DNA may be used as DNA probes in assays to
XX CC detect and quantitate the presence of similar DNAs in samples, and which
XX CC patients may need restorative therapy. The NOVX protein may also be used
XX CC as antigens in the production of antibodies (Abs) against NOVX and in
XX CC assays to identify modulators of NOVX expression and activity. The anti-
XX CC NOVX Abs and antagonist are used to down regulate expression and
XX CC activity. The anti-NOVX Abs are used for detecting the presence of NOVX
XX CC in samples. Disorders that may be prevented, diagnosed and/or treated
XX CC vary depending on the NOVX protein. NOV1, NOV3, NOV5, NOV7, NOV9-11 and
XX CC NOV13-16 have casein kinase II phosphorylation sites characteristic of
XX CC serine/threonine kinases, and are used to treat kinase-related disorders
XX CC (e.g. Peutz-Jeghers syndrome, cellular proliferation and contraception).
XX CC NOV2-3, NOV6 and NOV8 are homologous to the epidermal growth factor (EGF)
XX CC -like super family and are involved in, e.g. regulation of cell
XX CC development, apoptosis, cell adhesion, growth migration, cell structure
XX CC and motility and protein management, and are used to treat cancers,
XX CC inflammatory disorders, immune disorders and cellular adhesion disorders.
XX CC NOV6-10 are homologous to EGF-like fibrillin proteins and are used to
XX CC treat cardiovascular disease e.g. hypertrophic cardiomyopathy, long-QT
XX CC syndrome and marfan syndrome
XX SQ Sequence 669 AA;
Query March 26.3%; Score 93; DB 4; Length 669;
Best Local Similarity 100.0%; Pred. No. 1.5e-82;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MTGPGKHCKECSHVGDGLNCEPEQLPDRCLQNGQCHADAKCVDLHFQDTTVGVFHL 60
Db 566 MTGPGKHCKECSHVGDGLNCEPEQLPDRCLQNGQCHADAKCVDLHFQDTTVGVFHL 625
Qy 61 RSPGLQYKLTDFKAREACANEAATMATYNQLSY 93
Db 626 RSPGLQYKLTDFKAREACANEAATMATYNQLSY 658

RESULT 8
ABJ10588
ID ABJ10588 standard; protein; 897 AA.
XX AC ABJ10588;
XX XX 28-NOV-2002 (first entry)
XX XX Human novel protein NOV1c SEQ ID NO: 211.
XX KW Human; NOVX; single nucleotide polymorphism; SNP; anti-HIV; cytostatic;
XX KW antiarteriosclerotic; antidiabetic; antiaschmatic; antiinflammatory;
XX KW haemostatic; hypotensive; neuroprotective; anorectic; nootropic;
XX KW antidepressant; immunosuppressive; antibacterial; antiparasitic;
XX KW virucide; tranquilizer; anticonvulsant; osteopathic; analgesic;
XX KW antiparkinsonian; dermatological; antiinfertility; cerebroprotective;
XX KW antiaddictive.
XX OS Homo sapiens.
XX XX
XX XX Key Location/Qualifiers
XX FT Misc-difference 221 /note= "optionally Cys depending on polymorphism present
XX FT in coding sequence"
XX FT Misc-difference 325 /note= "optionally Leu depending on polymorphism present
XX FT in coding sequence"
XX FT Misc-difference 416 /note= "optionally Tyr depending on polymorphism present
XX FT in coding sequence"
XX FT Misc-difference 832 /note= "optionally Gly depending on polymorphism present
XX FT in coding sequence"
XX EN WO200259315-A2.
XX XX
XX PD 01-AUG-2002.
XX XX 19-DEC-2001; 2001WO-US050076.
XX XX 19-DEC-2000; 2000US-0256619P.
XX PR 18-JAN-2001; 2001US-0262959P.
XX PR 28-FEB-2001; 2001US-0272408P.
XX PR 20-APR-2001; 2001US-0285189P.
XX PR 26-JUL-2001; 2001US-0308039P.
XX PR 09-AUG-2001; 2001US-0311266P.
XX XX (CURA-) CURAGEN CORP.
XX XX Shimkets RA, Patturajan M, Vernet CAM, Casman SJ, Malyaukar U;
XX XX Shenoy S, Spytek KA, Gangolli E, Miller C, Boldog F, Li L;
XX XX Taupier RJ, Kekuda R, Smithson G, Zerhusen BD, Liu X, Colman SD;
XX XX Tcherev V, Si J, Edinger S, Stone D, Sciore P, Millet I;
XX XX Rothenberg M;
XX XX WPI; 2002-666903/71.
XX DR N-PSDB; ABT08489.
XX XX
XX PT New isolated NOVX polypeptides and polynucleotides, useful for
XX PT preventing, diagnosing or treating NOVX-associated disorders e.g.
XX PT diabetes, Crohn's disease, atherosclerosis, cancer, Huntington's disease
XX PT or Alzheimer's disease.
XX PS Claim 54; Page 25; 363pp; English.
XX XX
XX CC The present invention provides the protein and coding sequences of
XX CC several novel human proteins, designated NOVX. These can be used in the
XX CC treatment of diseases such as cancers, Hodgkin's disease, Von Hippel-
XX CC Lindau syndrome, Alzheimer's disease, stroke, tuberculous sclerosis,
XX CC hypercalcaemia, Parkinson's disease, Huntington's disease, cerebral
XX CC palsy, epilepsy, Lesch-Nyhan syndrome, multiple sclerosis, ataxia
XX CC telangiectasia, leukodystrophies, addiction, anxiety, depression, pain,
XX CC obesity, Crohn's disease, osteoporosis, inflammatory bowel disease,

CC infertility, atherosclerosis, hypertension, scleroderma, haemophilia,
CC diabetes, pancreatitis, autoimmune disease, asthma, arthritis,
CC immunodeficiencies, HIV, viral, bacterial or parasitic infections, or
CC graft-versus-host disease. The present sequence is a protein of the
CC invention
XX
SQ Sequence 897 AA;
Query Match 26.3%; Score 93; DB 5; Length 897;
Best Local Similarity 100.0%; Pred. No. 2e-82;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTGPGKHCKCKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVFHL 60
Db 489 MTGPGKHCKCKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVFHL 548
QY 61 RSPGLQYKLTDFKAREACANEAAATMATYNQLSY 93
Db 549 RSPGLQYKLTDFKAREACANEAAATMATYNQLSY 581
RESULT 9
ABJ10587
ID ABJ10587 standard; protein; 2420 AA.
XX
AC ABJ10587;
XX
DT 28-NOV-2002 (first entry)
XX
DE Human novel protein NOV1b SEQ ID NO: 4.
XX
KW Human; NOVX; single nucleotide polymorphism; SNP; anti-HIV; cytostatic;
KW antiarteriosclerotic; antidiabetic; antiasthmatic; antiinflammatory;
KW haemostatic; hypotensive; neuroprotective; anorectic; nootropic;
KW antidepressant; immunosuppressive; antibacterial; antiparasitic;
KW virucide; tranquilizer; anticonvulsant; osteopathic; analgesic;
KW antiparkinsonian; dermatological; antiinfertility; cerebroprotective;
KW antiaddictive.
XX
OS Homo sapiens.
XX
XX WO200259315-A2.
XX
XX 01-AUG-2002.
XX
XX 19-DEC-2001; 2001WO-US050076.
XX
XX 19-DEC-2000; 2000US-0256619P.
XX 19-JAN-2001; 2001US-0262959P.
XX 28-FEB-2001; 2001US-0272408P.
XX 20-APR-2001; 2001US-0285189P.
XX 26-JUL-2001; 2001US-0308039P.
XX 09-AUG-2001; 2001US-0311266P.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Shimkets RA, Patturajan M, Vernet CAM, Casman SJ, Malyankar U;
XX Shenoy S, Spytek KA, Gangolli E, Miller C, Boldog F, Li L;
XX Taupier RJ, Kekuda R, Smithson G, Zerhusen BD, Liu X, Colman SD;
XX Tchernev V, Si J, Edinger S, Stone D, Sciore P, Millet I;
XX Rothenberg M;
XX
XX WPI; 2002-666903/71.
XX
XX N-PSDB; ABT08488.
XX
XX New isolated NOVX polypeptides and polynucleotides, useful for
XX preventing, diagnosing or treating NOVX-associated disorders e.g.
XX diabetes, Crohn's disease, atherosclerosis, cancer, Huntington's disease
XX or Alzheimer's disease.
XX
XX Claim 1; Page 23-24; 363pp; English.
XX
XX The present invention provides the protein and coding sequences of

CC several novel human proteins, designated NOVX. These can be used in the
CC treatment of diseases such as cancers, Hodgkin's disease, Von Hippel-
CC Lindau syndrome, Alzheimer's disease, stroke, tuberosus sclerosis,
CC hypercalcaemia, Parkinson's disease, Huntington's disease, cerebral
CC palsy, epilepsy, Lesch-Nyhan syndrome, multiple sclerosis, ataxia
CC telangiectasia, leukodystrophies, addiction, anxiety, depression, pain,
CC obesity, Crohn's disease, osteoporosis, inflammatory bowel disease,
CC infertility, atherosclerosis, hypertension, scleroderma, haemophilia,
CC diabetes, pancreatitis, autoimmune disease, asthma, arthritis,
CC immunodeficiencies, HIV, viral, bacterial or parasitic infections, or
CC graft-versus-host disease. The present sequence is a protein of the
CC invention
XX
SQ Sequence 2420 AA;
Query Match 26.3%; Score 93; DB 5; Length 2420;
Best Local Similarity 100.0%; Pred. No. 4.5e-82;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTGPGKHCKCKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVFHL 60
Db 2250 MTGPGKHCKCKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTGVFHL 2309
QY 61 RSPGLQYKLTDFKAREACANEAAATMATYNQLSY 93
Db 2310 RSPGLQYKLTDFKAREACANEAAATMATYNQLSY 2342
RESULT 10
ABJ10586
ID ABJ10586 standard; protein; 2675 AA.
XX
AC ABJ10586;
XX
DT 28-NOV-2002 (first entry)
XX
DE Human novel protein NOV1a SEQ ID NO: 2.
XX
KW Human; NOVX; single nucleotide polymorphism; SNP; anti-HIV; cytostatic;
KW antiarteriosclerotic; antidiabetic; antiasthmatic; antiinflammatory;
KW haemostatic; hypotensive; neuroprotective; anorectic; nootropic;
KW antidepressant; immunosuppressive; antibacterial; antiparasitic;
KW virucide; tranquilizer; anticonvulsant; osteopathic; analgesic;
KW antiparkinsonian; dermatological; antiinfertility; cerebroprotective;
KW antiaddictive.
XX
OS Homo sapiens.
XX
XX WO200259315-A2.
XX
XX 01-AUG-2002.
XX
XX 19-DEC-2001; 2001WO-US050076.
XX
XX 19-DEC-2000; 2000US-0256619P.
XX 19-JAN-2001; 2001US-0262959P.
XX 28-FEB-2001; 2001US-0272408P.
XX 20-APR-2001; 2001US-0285189P.
XX 26-JUL-2001; 2001US-0308039P.
XX 09-AUG-2001; 2001US-0311266P.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Shimkets RA, Patturajan M, Vernet CAM, Casman SJ, Malyankar U;
XX Shenoy S, Spytek KA, Gangolli E, Miller C, Boldog F, Li L;
XX Taupier RJ, Kekuda R, Smithson G, Zerhusen BD, Liu X, Colman SD;
XX Tchernev V, Si J, Edinger S, Stone D, Sciore P, Millet I;
XX Rothenberg M;
XX
XX WPI; 2002-666903/71.
XX
XX N-PSDB; ABT08487.
XX
XX New isolated NOVX polypeptides and polynucleotides, useful for

PT preventing, diagnosing or treating NOVX-associated disorders e.g.
PT diabetes, Crohn's disease, atherosclerosis, cancer, Huntington's disease
XX or Alzheimer's disease.
XX
PS Claim 1; Page 20; 363pp; English.
XX

CC The present invention provides the protein and coding sequences of
CC several novel human proteins, designated NOVX. These can be used in the
CC treatment of diseases such as cancers, Hodgkin's disease, Von Hippel-
CC Lindau syndrome, Alzheimer's disease, stroke, tuberosus sclerosis,
CC hypercalcaemia, Parkinson's disease, Huntington's disease, cerebral
CC palsy, epilepsy, Leisch-Nyran syndrome, multiple sclerosis, ataxia
CC telangiectasia, leukodystrophies, addiction, anxiety, depression, pain,
CC obesity, Crohn's disease, osteoporosis, inflammatory bowel disease,
CC infertility, atherosclerosis, hypertension, scleroderma, haemophilia,
CC diabetes, pancreatitis, autoimmune disease, asthma, arthritis, or
CC immunodeficiencies, HIV, viral, bacterial or parasitic infections, or
CC graft-versus-host disease. The present sequence is a protein of the
CC invention
XX
XX Sequence 2675 AA;

Query Match 26.3%; Score 93; DB 5; Length 2675;
Best Local Similarity 100.0%; Pred. No. 4.9e-82;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTGPGKHCKCKSHVYGDGLNCEPEQLPIORCLQDNGQCHADAKCVLHFDQTTVGVFHL 60
Db 2303 MTGPGKHCKCKSHVYGDGLNCEPEQLPIORCLQDNGQCHADAKCVLHFDQTTVGVFHL 2362
QY 61 RSPILGQYKLTDFDKAREACANEATMATYNQLSY 93
Db 2363 RSPILGQYKLTDFDKAREACANEATMATYNQLSY 2395

RESULT 11
AAB42164
ID AAB42164 standard; protein; 330 AA.
XX
AC AAB42164;

XX
XX
XX 08-FEB-2001 (first entry)
XX
XX Human ORFX ORF1928 polypeptide sequence SEQ ID NO:3856.

XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnery; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antithumatic; antithyroid;
KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive.
XX
XX Homo sapiens.
XX
XX WO200058473-A2.
XX
XX 05-OCT-2000.
XX
XX 31-MAR-2000; 2000WO-US008621.
XX
XX 31-MAR-1999; 99US-0127607P.
XX
XX 02-APR-1999; 99US-0127636P.
XX
XX 05-APR-1999; 99US-0127728P.
XX
XX 30-MAR-2000; 2000US-00540763.

PA (CURA-) CURAGEN CORP.
XX
XX Shimkets RA, Leach M;
XX
XX WPI; 2000-602362/57.
XX
XX N-PSDB; AAC76373.

XX Novel nucleic acids and peptides derived from open reading frame X,
XX useful for treating e.g. cancers, proliferative disorders,
XX neurodegenerative disorders and cardiovascular disease.

XX Claim 11; Page 3007-3008; 5507pp; English.

XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
XX which represent the human ORFX open reading frames 1 to 3161. The ORFX
XX sequences have activities such as: cytostatic; hepatotropic; vulnery;
XX antipsoriatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;
XX anticonvulsant; antiarthritic; immunosuppressant; antidiabetic; hypotensive;
XX cardiant; thrombolytic; coagulant; vasotropic; antithumatic; antithyroid;
XX dermatological; immunosuppressive; antiinflammatory; antibacterial;
XX antiviral; antifungal; antithumatic; antithyroid; and antianemic. The
XX sequences can be used for determining the presence of or predisposition
XX to, or preventing or treating pathological conditions associated with an
XX ORFX-associated disorder. The nucleic acids can be used to express ORFX
XX proteins in gene therapy vectors. The proteins and nucleic acids may be
XX used to treat cancers, proliferative disorders, neurodegenerative
XX disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
XX diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
XX storage, systemic lupus erythematosus, severe combined immunodeficiency
XX (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
XX disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
XX cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
XX enhance coagulation; to inhibit thrombosis; and as a contraceptive

XX Sequence 330 AA;

Query Match 24.1%; Score 85; DB 3; Length 330;
Best Local Similarity 100.0%; Pred. No. 6.4e-75;
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 SGNLQVIMSPSLTNFLTEVLAYSNSRARGFLHLDLSIRGTLFVFPNSGLGENET 228
Db 89 SGNLQVIMSPSLTNFLTEVLAYSNSRARGFLHLDLSIRGTLFVFPNSGLGENET 148
QY 229 LSGRDIEHLANVSMFFYNDLVNGT 253
Db 149 LSGRDIEHLANVSMFFYNDLVNGT 173

RESULT 12
AAB83358
ID AAB83358 standard; protein; 315 AA.

XX
XX AAB83358;
XX
XX 26-MAR-2002 (first entry)
XX
XX NOV2 protein sequence.

XX NOV; Cytostatic; contraceptive; antiinflammatory; immunomodulatory;
KW cardiovascular; casein kinase II phosphorylation site; contraception;
KW serine/threonine kinase; Peutz-Jeghers syndrome; cellular proliferation;
KW epidermal growth factor; cell development; apoptosis; cell adhesion;
KW growth migration; cell structure; motility; cancer; immune disorder;
KW inflammatory disorder; cellular adhesion disorder; long-OT syndrome;
KW cardiovascular disease; hypertrophic cardiomyopathy; marfan syndrome;
KW therapy; NOV1; NOV2; NOV3; NOV4; NOV5; NOV6; NOV7; NOV8; NOV9; NOV10;
KW NOV11; NOV12; NOV13; NOV14; NOV15; NOV16.
XX
XX Unidentified.
XX
XX WO200136638-A2.

XX

PD 25-MAY-2001.
 XX 17-NOV-2000; 2000WO-US031543.
 PF 19-NOV-1999; 99US-0166336P.
 PR 29-NOV-1999; 99US-0167785P.
 PR 08-MAR-2000; 2000US-0187844P.
 PR 16-NOV-2000; 2000US-00715417.
 XX (CURA-) CURAGEN CORP.
 PA Shimkets RA, Lichenstein H, Vernet C, Fernandes E;
 XX WPI; 2001-648134/74.
 XX N-PSDB; AAF87113.
 DR Novel human polypeptides and the nucleic acids that encode them useful
 PT for preventing, diagnosing and treating e.g. cancer, inflammation and
 PT immune disorders.
 XX Claim 1; Page 10-13; 141pp; English.
 XX This sequence is the NOV2 protein. The invention relates to the NOV1-
 CC NOV6 proteins, and their coding sequences. The proteins have Cytostatic;
 CC contraceptive; antiinflammatory; immunomodulatory; and cardiovascular
 CC activities. The sequences may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate NOVX expression. They
 CC may be used to treat disorders associated with decreased expression by
 CC rectifying mutations or deletions in a patient's genome that affect the
 CC activity of protein by expressing inactive proteins or to supplement the
 CC patients own production of protein. They are used to produce NOVX
 CC proteins, by inserting the nucleic acid into a cell and culturing it to
 CC express the protein. The DNA may be used as DNA probes in assays to
 CC detect and quantitate the presence of similar DNAs in samples, and which
 CC patients may need restorative therapy. The NOVX protein may also be used
 CC as antigens in the production of antibodies (Abs) against NOVX and in
 CC assays to identify modulators of NOVX expression and activity. The anti-
 CC NOVX Abs and antagonist are used to down regulate expression and
 CC activity. The anti-NOVX Abs are used for detecting the presence of NOVX
 CC in samples. Disorders that may be prevented, diagnosed and/or treated
 CC vary depending on the NOVX protein. NOV1, NOV3, NOV5, NOV7, NOV9-11 and
 CC NOV13-16 have casein kinase II phosphorylation sites characteristic of
 CC serine/threonine kinases, and are used to treat kinase-related disorders
 CC (e.g. Peutz-Jeghers syndrome, cellular proliferation and contraception).
 CC NOV2-3, NOV6 and NOV8 are homologous to the epidermal growth factor (EGF)
 CC -like super family and are involved in, e.g. regulation of cell
 CC development, apoptosis, cell adhesion, growth migration, cell structure
 CC and motility and protein management, and are used to treat cancers,
 CC inflammatory disorders, immune disorders and cellular adhesion disorders.
 CC NOV6-10 are homologous to EGF-like fibrillin proteins and are used to
 CC treat cardiovascular disease e.g. hypertrophic cardiomyopathy, long-QT
 CC syndrome and marfan syndrome
 XX Sequence 315 AA;
 SQ Query Match 16.1%; Score 57; DB 4; Length 315;
 Best Local Similarity 100.0%; Pred. No. 2.3e-47;
 Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 QKAKYHLCAGWLETGRVAYPTAFASQNGSGVVGIVDYGRPNKSEMDVFCYRMK 151
 DB 239 QKAKYHLCAGWLETGRVAYPTAFASQNGSGVVGIVDYGRPNKSEMDVFCYRMK 295
 RESULT 13
 ABG72500
 ID ABG72500 standard; protein; 93 AA.
 XX AC ABG72500;
 XX 19-FEB-2003 (first entry)
 DT Human 190kDa Hyaluronan receptor for endocytosis (HARE) LINK domain.
 DE

XX Hyaluronan receptor for endocytosis; HARE; hyaluronan; HA; CD;
 KW chondroitin molecule; chondroitin sulphate; CDS; targeting; metastasis;
 KW tumour; gene therapy; human; LINK domain.
 XX Homo sapiens.
 XX WO200286093-A2.
 XX 31-OCT-2002.
 XX 25-APR-2002; 2002WO-US013209.
 XX 25-APR-2001; 2001US-00842930.
 PR 25-APR-2001; 2001US-0286468P.
 XX (WEIG/) WEIGEL P H.
 PA (WEIG/) WEIGEL J A.
 XX Weigel PH, Weigel JA;
 XX WPI; 2003-093126/08.
 XX Targeting compounds e.g. chemotherapeutic agent to cell of subject
 PT expressing functional active hyaluronan receptor for endocytosis of HARE,
 PT by using hyaluronan, HARE, and/or monoclonal antibody that binds to HARE
 PT epitope.
 XX Example; Page 77; 167pp; English.
 XX The invention describes a method of targeting a compound to a cell or
 CC tissue of an individual expressing a functionally active hyaluronan (HA)
 CC receptor for endocytosis (HARE) or a cell that does not express
 CC functionally active HARE. The method involves using HA molecule, a
 CC chondroitin molecule (CD), a chondroitin sulphate (CDS) molecule, and/or
 CC a monoclonal antibody raised against a HA-binding domain of HARE. The
 CC method is useful for targeting a compound, preferably a chemotherapeutic
 CC agent or a radioisotope to cell of an individual, especially a human,
 CC expressing HARE on its surface (e.g. gene therapy). Also described is a
 CC method useful for preventing interaction between a cell having at least
 CC one of a HA coat, a CD coat and CDS coat and a cell expressing HARE on
 CC its surface. This second method is useful for preventing metastasis by
 CC preventing interaction between tumour cells having HA, CD or CDS coat and
 CC non-tumour cells expressing HARE on its surface. The invention also
 CC describes a method useful for detecting the presence of HA, CD and CDS in
 CC a biological fluid. This sequence represents the LINK domain of the human
 CC 190kDa Hyaluronan receptor for endocytosis (HARE) that may be present in
 CC other HARE-like proteins
 XX SQ Sequence 93 AA;
 Query Match 15.3%; Score 54; DB 6; Length 93;
 Best Local Similarity 100.0%; Pred. No. 7.6e-45;
 Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 95 QKAKYHLCAGWLETGRVAYPTAFASQNGSGVVGIVDYGRPNKSEMDVFCY 148
 DB 40 QKAKYHLCAGWLETGRVAYPTAFASQNGSGVVGIVDYGRPNKSEMDVFCY 93
 RESULT 14
 AAM47675
 ID AAM47675 standard; protein; 1431 AA.
 XX AC AAM47675;
 XX 22-FEB-2002 (first entry)
 DT Rat Hyaluronic Acid Receptor for Endocytosis, HARE.
 DE HARE; Hyaluronic Acid Receptor for Endocytosis; hyaluronan; chondroitin;
 KW chondroitin sulphate; extracellular matrix; cartilage; skin;
 KW vitreous humour; endocytic receptor; glycosaminoglycan; rat.

PN WO200118014-A1.
 XX 15-MAR-2001.
 XX 30-AUG-2000; 2000WO-US023794.
 XX 03-SEP-1999; 99US-0152296P.
 XX 06-OCT-1999; 99US-0158003P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Roschke V;
 XX WPI; 2001-235186/24.
 XX N-PSDB; AAS00854.
 XX Twenty nine nucleic acid molecules encoding human cancer associated
 PT proteins, useful in the prevention, treatment and diagnosis of cancer,
 PT immune disorders, cardiovascular disorders and neurological diseases.
 XX Disclosure; Page 403; 427pp; English.
 XX The sequence represents a novel Human cancer related protein. The
 CC polynucleotides and polypeptides are useful for preventing, treating or
 CC ameliorating a medical condition in e.g. humans, mice, rabbits, goats,
 CC horses, cats, dogs, chickens or sheep. The polypeptides can also be used
 CC as a food additive or preservative to increase or decrease storage
 CC capabilities. The polynucleotide are useful for chromosome
 CC identification. The nucleic acids, protein, antibodies, agonists and
 CC antagonists are useful in the diagnosis, treatment and prevention of
 CC cancer (e.g. cancers of the adrenal gland, bone, bone marrow, breast,
 CC gastrointestinal tract, liver, lung, or urogenital), immune disorders
 CC (e.g. Addison's disease, allergies, autoimmune haemolytic anaemia,
 CC autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple
 CC sclerosis, rheumatoid arthritis and ulcerative colitis, acquired
 CC immunodeficiency syndrome, AIDS), cardiovascular disorders such as
 CC myocardial ischaemias, wound healing, neurological diseases (e.g.
 CC Parkinson's disease, Alzheimer's disease, cerebral anoxia and epilepsy)
 CC and infectious diseases such as viral, bacterial, fungal and parasitic
 CC infections. Numerous examples of each type of disorder are given in the
 CC specification
 XX Sequence 106 AA;
 SQ
 Query Match 5.7%; Score 20; DB 4; Length 106;
 Best Local Similarity 100.0%; Pred. No. 2.6e-11;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 326 IILVTGAVAAAYSYFRINR 345
 Db 23 IILVTGAVAAAYSYFRINR 42
 RESULT 17
 ABG72512
 ID ABG72512 standard; protein; 18 AA.
 AC ABG72512;
 XX
 XX 19-FEB-2003 (first entry)
 DT
 DE Human 190kDa Hyaluronan receptor for endocytosis putative motif #12.
 XX Hyaluronan receptor for endocytosis; HARE; hyaluronan; HA; CD;
 KW chondroitin molecule; chondroitin sulphate; targeting; metastasis;
 KW tumour; gene therapy; human; motif.
 XX Homo sapiens.
 OS
 XX WO200286093-A2.
 XX 31-OCT-2002.
 XX

PF 25-APR-2002; 2002WO-US013209.
 XX 25-APR-2001; 2001US-00842930.
 PR 25-APR-2001; 2001US-0286468P.
 XX (WEIG/) WEIGEL P H.
 PA (WEIG/) WEIGEL J A.
 XX Weigel PH, Weigel JA;
 PI WPI; 2003-093126/08.
 XX Targeting compounds e.g. chemotherapeutic agent to cell of subject
 PT expressing functional active hyaluronan receptor for endocytosis of HARE,
 PT by using hyaluronan, HARE, and/or monoclonal antibody that binds to HARE
 PT epitope.
 XX Claim 23; Page 77; 167pp; English.
 XX The invention describes a method of targeting a compound to a cell or
 CC tissue of an individual expressing a functionally active hyaluronan (HA)
 CC receptor for endocytosis (HARE) or a cell that does not express
 CC functionally active HARE. The method involves using HA molecule, a
 CC chondroitin molecule (CD), a chondroitin sulphate (CDS) molecule, and/or
 CC a monoclonal antibody raised against a HA-binding domain of HARE. The
 CC method is useful for targeting a compound, preferably a chemotherapeutic
 CC agent or a radioisotope to cell of an individual, especially a human,
 CC expressing HARE on its surface (e.g. gene therapy). Also described is a
 CC method useful for preventing interaction between a cell having at least
 CC one of a HA coat, a CD coat and CDS coat and a cell expressing HARE on
 CC its surface. This second method is useful for preventing metastasis by
 CC preventing interaction between tumour cells having HA, CD or CDS coat and
 CC non-tumour cells expressing HARE on its surface. The invention also
 CC describes a method useful for detecting the presence of HA, CD and CDS in
 CC a biological fluid. This sequence represents a putative motif from the
 CC human 190kDa Hyaluronan receptor for endocytosis (HARE) that may be
 XX present in other HARE-like proteins
 SQ Sequence 18 AA;
 Query Match 5.1%; Score 18; DB 6; Length 18;
 Best Local Similarity 100.0%; Pred. No. 5.5e-10;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 TGPGRKCKCKSHYVGDG 19
 Db 1 TGPGRKCKCKSHYVGDG 18
 RESULT 18
 AA47702
 ID AA47702 standard; peptide; 15 AA.
 XX AA47702;
 XX 22-FEB-2002 (first entry)
 DT
 DE Human HARE peptide fragment #6.
 XX HARE; Hyaluronic Acid Receptor for Endocytosis; hyaluronan; chondroitin;
 KW chondroitin sulphate; extracellular matrix; cartilage; skin; rat;
 KW vitreous humour; endocytic receptor; glycosaminoglycan.
 XX Homo sapiens.
 OS
 XX WO200181544-A2.
 XX 01-NOV-2001.
 XX 25-APR-2001; 2001WO-US013403.
 XX 25-APR-2000; 2000US-0199538P.
 PR 02-NOV-2000; 2000US-0245320P.
 PR

XX (WEIG/) WEIGEL P A.
 PA (ZHOU/) ZHOU B.
 PA (WEIG/) WEIGEL J A.
 XX Weigel PA, Zhou B, Weigel JA;
 XX WPI; 2002-049271/06.
 DR New mammalian hyaluronic acid receptor for endocytosis, useful e.g. for
 PT identifying agents that inhibit binding to hyaluronic acid, and related
 PT nucleic acid.
 XX
 XX Example; Page 95; 263pp; English.
 PS
 CC The present invention relates to sequences for rat and human HARE
 CC (Hyaluronic Acid Receptor for Endocytosis, ABA04648, ABA04662, AAM47675
 CC and AAM47684). HARE can bind specifically to at least one of hyaluronic
 CC acid (HA, also known as hyaluronan), chondroitin (C) or chondroitin
 CC sulphate (CS). HA is an extracellular matrix component of all tissues, in
 CC particular cartilage, skin and vitreous humour. HARE is the endocytic
 CC receptor responsible for removing HA and other glycosaminoglycans from
 CC the circulation. The present sequence is a peptide fragment of HARE,
 CC which was used in an example from the present invention
 XX
 XX Sequence 15 AA;
 SQ
 Query Match 4.2%; Score 15; DB 5; Length 15;
 Best Local Similarity 100.0%; Pred. No. 4.3e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 61 RSPGLGQYKLTFFDKAR 75
 DB 1 RSPGLGQYKLTFFDKAR 15
 RESULT 19
 AAM47690
 ID AAM47690 standard; peptide; 13 AA.
 AC AAM47690;
 XX
 XX 22-FEB-2002 (first entry)
 DT
 DE Human HARE peptide fragment PR 1825 #2.
 DE
 KW HARE; Hyaluronic Acid Receptor for Endocytosis; hyaluronan; chondroitin;
 KW chondroitin sulphate; extracellular matrix; cartilage; skin; human; rat;
 KW vitreous humour; endocytic receptor; glycosaminoglycan.
 XX
 OS Homo sapiens.
 OS
 PN WO200181544-A2.
 PN
 PD 01-NOV-2001.
 PD
 XX 25-APR-2001; 2001WO-US013403.
 PF
 XX 25-APR-2000; 2000US-0199538P.
 PR
 XX 02-NOV-2000; 2000US-0245320P.
 PR
 XX (WEIG/) WEIGEL P A.
 PA (ZHOU/) ZHOU B.
 PA (WEIG/) WEIGEL J A.
 XX Weigel PA, Zhou B, Weigel JA;
 XX WPI; 2002-049271/06.
 DR New mammalian hyaluronic acid receptor for endocytosis, useful e.g. for
 PT identifying agents that inhibit binding to hyaluronic acid, and related
 PT nucleic acid.
 PT
 XX

PS Example; Page 94; 263pp; English.
 XX
 CC The present invention relates to sequences for rat and human HARE
 CC (Hyaluronic Acid Receptor for Endocytosis, ABA04648, ABA04662, AAM47675
 CC and AAM47684). HARE can bind specifically to at least one of hyaluronic
 CC acid (HA, also known as hyaluronan), chondroitin (C) or chondroitin
 CC sulphate (CS). HA is an extracellular matrix component of all tissues, in
 CC particular cartilage, skin and vitreous humour. HARE is the endocytic
 CC receptor responsible for removing HA and other glycosaminoglycans from
 CC the circulation. The present sequence is a peptide fragment of HARE,
 CC which was used in an example from the present invention
 XX
 XX Sequence 13 AA;
 SQ
 Query Match 3.7%; Score 13; DB 5; Length 13;
 Best Local Similarity 100.0%; Pred. No. 3.5e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 213 GTLFVPQNSGLGE 225
 DB 1 GTLFVPQNSGLGE 13
 RESULT 20
 ABG72513
 ID ABG72513 standard; protein; 13 AA.
 AC ABG72513;
 XX
 XX 19-FEB-2003 (first entry)
 DT
 DE Human 190kDa Hyaluronan receptor for endocytosis putative motif #13.
 DE
 KW Hyaluronan receptor for endocytosis; HARE; hyaluronan; HA; CD;
 KW chondroitin molecule; chondroitin sulphate; CDS; targeting; metastasis;
 KW tumour; gene therapy; human; motif.
 XX
 OS Homo sapiens.
 OS
 PN WO200286093-A2.
 PN
 PD 31-OCT-2002.
 PD
 XX 25-APR-2002; 2002WO-US013209.
 PF
 XX 25-APR-2001; 2001US-00842930.
 PR
 XX 25-APR-2001; 2001US-0286468P.
 PR
 XX (WEIG/) WEIGEL P H.
 PA (WEIG/) WEIGEL J A.
 XX Weigel PH, Weigel JA;
 XX WPI; 2003-093126/08.
 DR
 XX Targeting compounds e.g. chemotherapeutic agent to cell of subject
 PT expressing functional active hyaluronan receptor for endocytosis of HARE,
 PT by using hyaluronan, HARE, and/or monoclonal antibody that binds to HARE
 PT epitope.
 XX
 XX Claim 23; Page 77; 167pp; English.
 PS
 CC The invention describes a method of targeting a compound to a cell or
 CC tissue of an individual expressing a functionally active hyaluronan (HA)
 CC receptor for endocytosis (HARE) or a cell that does not express
 CC functionally active HARE. The method involves using HA molecule, a
 CC chondroitin molecule (CD), a chondroitin sulphate (CDS) molecule, and/or
 CC a monoclonal antibody raised against a HA-binding domain of HARE. The
 CC method is useful for targeting a compound, preferably a chemotherapeutic
 CC agent or a radioisotope to cell of an individual, especially a human,
 CC expressing HARE on its surface (e.g. gene therapy). Also described is a
 CC method useful for preventing interaction between a cell having at least
 CC one of a HA coat, a CD coat and CDS coat and a cell expressing HARE on

CC its surface. This second method is useful for preventing metastasis by
 CC preventing interaction between tumour cells having HA, CD or CDS coat and
 CC non-tumour cells expressing HARE on its surface. The invention also
 CC describes a method useful for detecting the presence of HA, CD and CDS in
 CC a biological fluid. This sequence represents a putative motif from the
 CC human 190Kda Hyaluronan receptor for endocytosis (HARE) that may be
 CC present in other HARE-like proteins
 XX

SQ Sequence 13 AA;

Query Match 3.7%; Score 13; DB 6; Length 13;
 Best Local Similarity 100.0%; Pred. No. 3.5e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 28 PIDRCLQDNGQCH 40
 |||||
 Db 1 PIDRCLQDNGQCH 13

Search completed: April 1, 2004, 09:56:56
 Job time : 59 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 1, 2004, 09:56:50 ; Search time 3238 Seconds
(without alignments)
3255.513 Million cell updates/sec

Title: US-09-466-778B-11
Perfect score: 1865
Sequence: 1 MTGPGKHCKSHYVDGL.....ALAAYSYFIRNKRKTIGFXHF 353

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09466778/runat_01042004_095642_10592/app_query.fasta_1.519
-DB=EST -QFMT=fastcap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOFCU=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09466778 @CIGN 1.1 3437 @runat_01042004_095642_10592 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCAPOP=6
-FCAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:

- 1: em estba:*
- 2: em esthum:*
- 3: em estin:*
- 4: em estmu:*
- 5: em estov:*
- 6: em estpl:*
- 7: em estro:*
- 8: em htc:*
- 9: gb est1:*
- 10: gb est2:*
- 11: gb htc:*
- 12: gb est3:*
- 13: gb est4:*
- 14: gb est5:*
- 15: em estfun:*
- 16: em estom:*
- 17: em gss hum:*
- 18: em gss inv:*
- 19: em gss pln:*
- 20: em gss vrt:*
- 21: em gss fun:*
- 22: em gss mam:*
- 23: em gss mus:*
- 24: em gss pro:*
- 25: em gss rod:*
- 26: em gss phg:*
- 27: em gss vrl:*
- 28: gb gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1487	79.7	3085	11	AK034522	AK034522 Mus muscu
2	963	51.6	819	13	BY717115	BY717115 BX717115
3	718	38.5	694	14	CA376384	CA376384 654727 NC
4	703	37.7	683	12	BJ523552	BJ523552 BJ523552
5	661.5	35.5	3025	11	BC049247	BC049247 Mus muscu
6	638	34.2	1122	14	CK028027	CK028027 AGENCOURT
7	604.5	32.4	722	13	BY739778	BY739778 BX739778
8	590	31.6	733	12	BI832486	BI832486 603082278
9	575.5	30.9	2368	11	BC019712	BC019712 Mus muscu
10	557.5	29.9	459	14	T47504	T47504 Yb14f01.r1
11	552	29.6	3454	11	AK028441	AK028441 Mus muscu
12	504.5	27.1	1201	13	BX446538	BX446538 BX446538
13	462	24.8	572	14	CB268774	CB268774 1007680 H
14	441.5	23.7	865	13	BY705197	BY705197 BX705197
15	441	23.6	719	14	CA342417	CA342417 672284 NC
16	424.5	22.8	761	13	BX861168	BX861168 BX861168
17	418.5	22.4	878	14	CD359001	CD359001 AGENCOURT
18	384.5	20.6	690	10	BF342700	BF342700 602032375
19	384	20.6	962	10	BF300451	BF300451 602032375
20	375	20.1	570	10	BE233224	BE233224 139315 MA
21	356	19.1	557	12	BM258620	BM258620 523466 MA
22	353.5	19.0	798	14	CK030769	CK030769 AGENCOURT
23	350	18.8	721	12	BI219897	BI219897 602936347
24	345.5	18.5	646	12	BM767207	BM767207 K-EST0049
25	344	18.4	895	29	CNS03DRC	AL238989 Tetraodon
26	343.5	18.4	589	12	BM769767	BM769767 K-EST0053
27	339	18.2	522	13	BX866658	BX866658 BX866658
28	335	18.0	388	10	AW435659	AW435659 74622 MAR
29	334	17.9	1253	13	BQ884322	BQ884322 AGENCOURT
30	333.5	17.9	725	14	CA362129	CA362129 636116 NC
31	329.5	17.7	564	12	BM181159	BM181159 f94h10.Y
32	329.5	17.7	567	12	BM181146	BM181146 f94h10.Y
33	320.5	17.2	928	12	BI906024	BI906024 603062355
34	316	16.9	804	12	BI144649	BI144649 602909939
35	313.5	16.8	467	12	BI132408	BI132408 AR030605L
36	309.5	16.6	971	29	CNS03XY8	AL265625 Tetraodon
37	304.5	16.3	429	10	BF870695	BF870695 IL3-ET011
38	296.5	15.9	439	12	BM694227	BM694227 UI-E-C11-
39	293.5	15.7	482	12	BM769766	BM769766 K-EST0053
40	289.5	15.5	704	14	CA364659	CA364659 639611 NC
41	285.5	15.3	520	12	BG712246	BG712246 pgl1n.pk0
42	282	15.1	603	13	BQ131025	BQ131025 f247905.Y
43	282	15.1	604	10	AW595426	AW595426 f247902.Y
44	281	15.1	789	12	BG935766	BG935766 SSI-0071
45	280	15.0	606	14	CA347334	CA347334 678437 NC

ALIGNMENTS

RESULT 1
AK034522
LOCUS
DEFINITION
AK034522 Mus musculus adult male diencephalon cDNA, RIKEN full-length enriched library, clone:9330210123 product:similar to CD44-LIKE PRECURSOR FELL (Homo sapiens), full insert sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
AK034522.1 GI:26330004
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

1 Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

2 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

3 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsuunai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Okazaki,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
Yoneda,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (KISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

4 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 695-690 (2001)
11076861

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

5 The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
11076861

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

6 (bases 1 to 3085)
Adachi,J., Aizawa,K., Akimura,T., Atakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Takeda,Y., Tanaka,I., Tomaru,A., Toyota,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

7 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 695-690 (2001)
11076861

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

8 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 695-690 (2001)
11076861

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

9 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 695-690 (2001)
11076861

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

10 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 695-690 (2001)
11076861

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

11 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 695-690 (2001)
11076861

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

12 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 695-690 (2001)
11076861

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

13 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 695-690 (2001)
11076861

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

14 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 695-690 (2001)
11076861

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

15 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 695-690 (2001)
11076861

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

16 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 695-690 (2001)
11076861

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

17 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 695-690 (2001)
11076861

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

18 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 695-690 (2001)
11076861

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

19 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 695-690 (2001)
11076861

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

20 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 695-690 (2001)
11076861

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

21 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 695-690 (2001)
11076861

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

22 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 695-690 (2001)
11076861

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

23 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 695-690 (2001)
11076861

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

24 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 695-690 (2001)
11076861

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

25 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 695-690 (2001)
11076861

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

26 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 695-690 (2001)
11076861

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

27 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 695-690 (2001)
11076861

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

28 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 695-690 (2001)
11076861

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

29 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 695-690 (2001)
11076861

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

30 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 695-690 (2001)
11076861

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

31 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 695-690 (2001)
11076861

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

32 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 695-690 (2001)
11076861

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

33 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 695-690 (2001)
11076861

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

34 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409,

```
Db 1816 AAGAAATGT--GCAACATTGTGGGATTTAGACTACGGAACCGAGACCAACAGAGT 1872
Qy 141 GluMetTrrpAspValPheCysTyrArgMetLysAspValAenCysThr***LysValGly 160
Db 1873 GAATGTGGGACGCTCTCTGTACAGATGAAGATGTAAGTAACTGCACTGCAAGCGCAGG 1932
Qy 161 TyrValGlyAspGlyPheSerTyrSerGlyAenLeuGlnValLeuMetSerPhePro 180
Db 1933 TACGTGGGCGATGCTTCTCTGCAATGGAACTGTCTGGCGGGACATTGAGCACCACTCACTAAT 1992
Qy 181 SerLeuThrAenPheLeuThrGluValLeuAlaTyrSerAenSerSerAlaAargGlyArg 200
Db 1993 TCATCTACGAATCTCTGACAGAGGTGCTGTTTTTTCCAGGAGCTCAGCCCAAGGCGCG 2052
Qy 201 AlaPheLeuGluHisLeuThrAspLeuSerIleAargGlyThrLeuPheValProGlnAsn 220
Db 2053 GCGTTTTTGAACACCTGACTGCTGCTCAATCAGTGGCACCTGTTGTGCCACAGAAC 2112
Qy 221 SerGlyLeuGlyGluAenGluThrLeuSerGlyAargAspIleGluHisLeuAlaAsn 240
Db 2113 AGTGGGCTACCGAAAAATAGAGGCTGTCTGGCGGGACATTGAGCACCACTCACTAAT 2172
Qy 241 ValSerMetPhePheTyrAsnAspLeuValAenGlyThrThrLeuGlnThrAargLeuGly 260
Db 2173 GTCACAGTCTCTTTTACGATGACCTTGTCAATGGTACCTGCTTGAAGACTAGGCTGGGA 2232
Qy 261 SerLysLeuLeuIleThrAspAargGlnAspProLeuHisProThrGluThrAargCysVal 280
Db 2233 AGCCAACCTGCTCATCCTCCAGCCAGGACCGACCTCCAC---CAAGAGCGCAGGTTCGTG 2289
Qy 281 AspGlyAargAspThrLeuGluThrPaspIleCysAlaSerAenGlyIleThrHisValIle 300
Db 2290 GATGGAAGAGCCATTCTGCACTGGGACATCATTTGCTCTTAACGGGGTTCTCATATCAT 2349
Qy 301 SerArg***LeuLysAlaProProAlaProValThrLeu***HisThrGlyLeuGly*** 320
Db 2350 TCTGAACCTTTTGAAGCTCTCTCCAGCGCCGACCGCTGCCACTCTGGCTTAGGAACA 2409
Qy 321 GlyIlePhe*****IleLeuValThrGlyAlaValAlaLeuAlaAlaTyrSerTyr 340
Db 2410 GGCATATTCGCGTGTCTGCTGTTACTGGTGGATGCTGCTGGCTGCTCTCTCTCTAC 2469
Qy 341 PheArgIleAenAargLysThrIleGlyPhe***HisPhe 353
Db 2470 TTCGGCTAAACAGAGAACCACTGGCTTCGGGGGTTT 2508

RESULT 2
BX717115 819 bp mRNA linear EST 18-NOV-2003
LOCUS BX717115 XGC-tadpole Silurana tropicalis cDNA clone TtpA026k09 5',
DEFINITION mRNA sequence.
ACCESSION BX717115 GI:38389795
VERSION BX717115
KEYWORDS EST.
SOURCE Silurana tropicalis (western clawed frog)
ORGANISM Silurana tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
Xenopodinae; Silurana.
1 (bases 1 to 819)
Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
Sanger Xenopus tropicalis EST project 2001 (11_2003)
Unpublished (2003)
Contact: Croning MDR
Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS SEQUENCE ID: TtpA026k09.p1kSP6
Sequencing primer: SP6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Nigel Garrett.
cDNA was oligo dt primed from Sug of poly A+ RNA from tadpole
```

embryos. EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the 5' end and NotI at the 3' end.
Vector: pCS107; Site 1: EcoRI; Site 2: NotI
Host: Escherichia coli DH10B.
FEATURES
Location/Qualifiers
source 1..819
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="TtpA026k09"
/dev_stage="tadpole (stage 35-40)"
/lab_host="E. coli DH10B"
/clone_lib="XGC-tadpole"
/notes="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA was oligo dt primed from Sug of poly A+ RNA from tadpole embryos. EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the 5' end and NotI at the 3' end"

ORIGIN

```
Alignment Scores:
Pred. No.: 5,81e-107 Length: 819
Score: 963.00 Matches: 179
Percent Similarity: 79.12% Conservative: 37
Best Local Similarity: 65.57% Mismatches: 51
Query Match: 13 Indels: 6
DB: 1 Gaps: 1

US-09-466-778b-11 (1-353) x BX717115 (1-819)

Qy 15 TyrValGlyAspGlyLeuAenCysGluProGluGlnLeuProIleAspAargCysLeuGln 34
Db 1 TATATTGGCATGGGTGGAGTTGTAAGTGAAGAGCTTCCAAATTAATGCTGCTGCTCAA 60
Qy 35 AspAsnGlyGlnCysHisAlaAspAlaLysCysValAspLeuHisPheGlnAspThrThr 54
Db 61 GACAAAGGGCAATGCCATGCTGATGCTATGTAATGATCTACACTTACAGATTCCACC 120
Qy 55 ValGlyValPheHisLeuAargSerProLeuGlyGlnTyrLysLeuThrPheAspLysAla 74
Db 121 GTCGGTGTGTTTCATTTTCGTTCTCCAAAGGGGTGTACAAATACAACTACAGCGAAGCC 180
Qy 75 ArgGluAlaCysAlaAenGluAlaAlaThrMetAlaThrTyrAenGlnLeuSerTyr*** 94
Db 181 ATGAAGCTTTGTAATGATGAGACTGCAACAATAGCAACTCAATCAGCTGTCTATATGCA 240
Qy 95 GlnLysAlaLysTyrHisLeuCysSerAlaGlyTrrpLeuGluThrGlyAargValAlaTyr 114
Db 241 CAACAGCGGGATTCCTCTCTGTCAGCTGGTGGCTGGATGGCTTAGAGTCGCATAT 300
Qy 115 ProThrAlaPheAlaSerGlnAenCysGlySerGlyValValGlyIleValAspTyrGly 134
Db 301 CCTACCACTTATTCCAAACCAAGACTGTGGATCAGGATTTGTTGGAATAGTACACTATGA 360
Qy 135 ProArgProAenLysSerGluMetTrrpAspValPheCysTyrAargMetLysAspValAsn 154
Db 361 CCCCGTGTCAACCTTAAGTGAAGCTGAGATGTTTTCTGCTACAGATTAAGAAGTGTGCA 420
Qy 155 CysThr***LysValGlyTyrValGlyAspGlyPheSerTyrSerGlyAenLeuGln 174
Db 421 TGTATTGCCAAACCTGGCTATGTGGAGATGGCTACACTTGTAAATGGAACCTGCTCAA 480
Qy 175 ValLeuMetSerPheProSerLeuThrAenPheLeuThrGluValLeuAlaTyrSerAsn 194
Db 481 GTCTTGACATCTTTTCCTTTCCTTCAAAATTTCTTACGGAATTTTGGTGTACTCTAAC 540
Qy 195 SerSerAlaargGlyAargAlaPheLeuGluHisLeuThrAspLeuSerIleAargGlyThr 214
Db 541 ACTTCAGTGAAGGGAAGAAATTTTCAACTACCTCACTAATGAAAATCAGACCTGTCTGGACGGGATATA 600
Qy 215 LeuPheValProGlnAenSerGlyLeuGlyGluAenGlnThrLeuSerGlyAargAspIle 234
Db 601 CTGTTTGGCCCAAGCAATGATGGACTTAATGAAAATCAGACCTGTCTGGACGGGATATA 660
```

```

QY 235 GluHisLeuAlaAsnValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThr 254
      |||...|||
Db 661 GAATATCATCTGGCAATGTCAGCATGTTCTTTGAGGACCTGAGCAATGGAACAACT 720
QY 255 LeuGlnThrArgLeuGlySerLysLeuLeuThrThrAspArgGlnAspProLeuHisPro 274
      |||...|||
Db 721 CTTCAACAGCAATAGGCCACAAAGTATTAAATTTCTTTTGAATGACCCAGCCAGCAAA 780
QY 275 -----ThrGluThrArgCysValAsp 281
      |||...|||
Db 781 ATAACCACTGATTCATGACTCCAAACCATGATGTTGGAT 819

RESULT 3
LOCUS CA376384 694 bp mRNA linear EST 06-NOV-2002
DEFINITION 654727 NCCOWA 1RT Oncorhynchus mykiss cdna clone IRT35N18_D G09 5',
      mRNA sequence.
ACCESSION CA376384
VERSION CA376384.1 GI:24694516
KEYWORDS EST.
SOURCE Oncorhynchus mykiss (rainbow trout)
ORGANISM Oncorhynchus mykiss
      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Actinopterygii; Neopterygii; Teleostei; Euteleostei;
      Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE 1 (bases 1 to 694)
      Rexroad, C.E. and Keele, J.W.
      Sequence analysis of a rainbow trout normalized cdna library
      Unpublished (2002)
COMMENT Contact: Rexroad CE
      USDA, ARS, National Center for Cool and Cold Water Aquaculture
      11876 Leetown Road, Kearneysville, WV 25430, USA
      Tel: 304 724 8340 x2129
      Fax: 304 725 0351
      Email: crexroad@cccwa.ars.usda.gov
      Single pass sequencing. Bases called with phred v0.020425, c and
      trimmed with the aid of the trim_alt option. Vector identified by
      cross_match v0.990329.
      Seq primer: AGCGGATACAAATTTTCACACAGGA.

FEATURES             Location/Qualifiers
     source           1..694
                     /organism="Oncorhynchus mykiss"
                     /mol_type="mRNA"
                     /db_xref="taxon:8022"
                     /clone="IRT35N18_D G09"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /clone_lib="NCCOWA 1RT"
                     /note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
                     Library made from pooled tissue from brain, gill, liver,
                     spleen, muscle, and kidney."

ORIGIN
Alignment Scores: 5.32e-77 Length: 694
Pred. No.: 718.00 Matches: 126
Score: 718.00 Mismatches: 45
Percent Similarity: 74.67% Conservative: 45
Best Local Similarity: 55.02% Mismatches: 58
Query Match: 38.50% Indels: 0
DB: 14 Gaps: 0

US-09-466-778b-11 (1-353) x CA376384 (1-694)

QY 25 GluGlnLeuProLeuAspArgCysLeuGlnAspAsnGlyGlnCysHisAlaAspAlaLys 44
      |||...|||
Db 4 AAGAGCTTCTTGACACCTGTTCTTCAGGAGAACGGGACGTGCCACTTCAGCGCCAG 63
QY 45 CysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeuArgSerProLeu 64
      |||...|||
Db 64 TGTACCGACCTTCACTACGAGGACAGCACAGTCGGTGTGTTCCACTTCGGGTCCCAACAG 123
QY 65 GlyGluTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsnGluAlaThr 84
      |||...|||

```

```

Db 124 GGACAGTACAACTGAACTATACATCGGCCCCAGGAGGCTCCACAGAGAGGGGGGAACC 183
QY 85 MetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHisLeuCysSerAla 104
      |||...|||
Db 184 ATCCGACGATACACACAGCTCTCTACGCACAGCAGGCTGGGTTCAACTGTGTGGCGGT 243
QY 105 GlyTrpLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSerGlnAsnCysGly 124
      |||...|||
Db 244 AGCTGGCTGGACCCAGGCGCAGAGTGGCTTACCCACCACCTACTCTCAACCCCAAGTGTGGC 303
QY 125 SerGlyValValGlyValLeuValAspTyrGlyProArgProAsnLysSerGluMetTrpAsp 144
      |||...|||
Db 304 TTTGGTACGTTGGGCGATCGTAGACTACGGCTTACGCAACACCTCAGTAGAGACCTGGGAC 363
QY 145 ValPheCysTyrArgMetLysAspValAsnCysThr***LysValGlyTyrValGlyAsp 164
      |||...|||
Db 364 ACCTTCTGCTACAGGCTCAAGAGGTGAAATGTGAATGAAGCAGAGGTATATAGGAGAT 423
QY 165 GlyPheSerTyrSerGlyAsnLeuGlnValLeuMetSerPheProSerLeuThrAsn 184
      |||...|||
Db 424 GGATACCTCTGCACTGGAAACCTCTCGCAGGTTCTCACTGAGAAGCCACCTCTCCAAC 483
QY 185 PheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArgAlaPheLeuGlu 204
      |||...|||
Db 484 TTCCTCTCTCAATCTGAACTACTCCAGATGTCCTGTCAGGAAACAGTTTATGATG 543
QY 205 HisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsnSerGlyLeuGly 224
      |||...|||
Db 544 CGCTCAGTAACCTGACCATCCAGTCCACTCTGTTGTACCTGATATATACCGGACTGTAC 603
QY 225 GluAsnGluThrLeuSerGlyArgAspIleGluHisLeuAlaAsnValSerMetPhe 244
      |||...|||
Db 604 CAGAACCCAGCACTGACCCACAGGACATGAGTACCATCTCTCAGAGGGCGGCGCTTG 663
QY 245 PheTyrAsnAspLeuValAsnGlyThr 253
      |||...|||
Db 664 GCCCTGAAAGACCTGACCAACGGCAGC 690

RESULT 4
LOCUS BJ523552 683 bp mRNA linear EST 09-AUG-2002
DEFINITION BJ523552 MF01SSB cdna Oryzias latipes cdna clone MF01SSB002F09 5',
      mRNA sequence.
ACCESSION BJ523552
VERSION BJ523552.1 GI:22182364
KEYWORDS EST.
SOURCE Oryzias latipes (Japanese medaka)
ORGANISM Oryzias latipes
      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
      Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
      Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
REFERENCE 1 (bases 1 to 683)
      Kohara, Y., Shin-i, T., Kimura, T., Narita, T., Jindo, T. and Takeda, H.
      Medaka EST Project in Takeda's lab
      Unpublished (2001)
      Contact: Tadasu Shin-i
      Center For Genetic Resource Information
      National Institute of Genetics
      1111 Yata, Mishima, Shizuoka 411-8540, Japan
      Tel: 81-559-81-6856
      Fax: 81-559-81-6855
      Email: tshini@genes.nig.ac.jp.
      Location/Qualifiers
           1..683
           /organism="Oryzias latipes"
           /mol_type="mRNA"
           /strain="Hd-rR"
           /db_xref="taxon:8090"
           /clone="MF01SSB002F09"
           /sex="mixture of female and male"
           /tissue_type="whole embryo"
           /dev_stage="segmentation stage 20 - 25"

FEATURES             source

```

/clone_lib="MF01SSB cdna"

ORIGIN

Alignment Scores:

Pred. No.: 3,59e-75 Length: 683
 Score: 703.00 Matches: 123
 Percent Similarity: 73.73% Conservative: 37
 Best Local Similarity: 56.68% Mismatches: 57
 Query Match: 37.69% Indels: 0
 DB: 12 Gaps: 0

US-09-466-778B-11 (1-353) x BJ523552 (1-683)

QY 1 MetThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeu 20
 Db 32 ATGACGCTCGGGAAAAAGAACTGGCCCTGTGAAGCACTACCTTGGAGATGGCGTG 91
 QY 21 AsnCysGluProGluGlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis 40
 Db 92 AGTGGCACTCATGCAAGATCCCGTGTAGCGCTGTCTCAGGCAATGGGGGTGCCAC 151
 QY 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60
 Db 152 TCAGATGCCAGATGTTCTGATCTCCACTTGAAGATAAAGTGTGTGTTCAATGTC 211
 QY 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80
 Db 212 CGCTCCAAACCGGTGAGTAACTACCTATGCTGACGCTCAGCATCGAGTTTGCATCATA 271
 QY 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr**GlnLysAlaLysTyrHis 100
 Db 272 GAGGAGGGTCCCTTGGCCAGCGCTCCAGCTTCCAGTCTCAGCAGGGGGGCTTGAAC 331
 QY 101 LeuCysSerAlaGlyTyrLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120
 Db 332 ATGTGTGCTGCTGCTGTTGGACCAAGCGAGTTGCATACCCCACTACCTACTCTPAAC 391
 QY 121 GlnAsnCysGlySerGlyValValGlyValAlaAspTyrGlyProArgProAsnLysSer 140
 Db 392 CCCAATGTGGCTTGGACATGGGGTATGTGACTACCGGTTCCGCAAAAACCTGAGC 451
 QY 141 GluMetTyrAspValPheCysTyrArgMetLysAspValAsnCysThr**LysValGly 160
 Db 452 GAGACCTGGGACGGCTTCTGTACCGGATGAAGATGTGAATGTCAATTAACCTGGT 511
 QY 161 TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuGlnValLeuMetSerPhePro 180
 Db 512 TATATCGAGACGGTGTGTGCTGTGTGTAACCTGATGCGAGTCTCTTAAGTCCAGCCC 571
 QY 181 SerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArg 200
 Db 572 AGCTTTCTAACTTCCTCACAAATCCTAAATTAATCTCCAGACAGCTCCNAGTCCGGCCG 631
 QY 201 AlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheVal 217
 Db 632 CAGTTGTCAACGCGCTCGGTGACCTCGAGTTCAGTTCAGTCCACCTGTGTGA 582

RESULT 5

BC049247
 LOCUS 3025 bp mRNA linear HTC 28-MAR-2003
 DEFINITION Mus musculus, similar to stabilin 1, clone IMAGE:5038209, mRNA.
 ACCESSION BC049247
 VERSION BC049247.1 GI:29351595
 KEYWORDS HTC.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 3025)
 STRAUSBERG.R.
 REFERENCE Direct Submission
 AUTHORS Submitted (21-MAR-2003) National Institutes of Health, Mammalian
 JOURNAL Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: mgc-help@nci.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)

DNA Sequencing by: National Institutes of Health Intramural

Sequencing Center (NISC),

Gaithersburg, Maryland;

Web site: <http://www.nisc.nih.gov/>

Contact: nisc_mgc@nci.nih.gov

Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,

Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,

Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,

Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,

Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,

McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,

Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,

Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 99 Row: 9 Column: 14
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis
 This clone has the following problem: frame shifted.

FEATURES

source

1..3025
 /location/Qualifiers
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="CZECH II"
 /db_xref="taxon:10090"
 /clone="IMAGE:5038209"
 /tissue_type="Mammary tumor metastasized to lung. Tumor
 arose spontaneously from a senescent normal mammary
 (clonal) outgrowth infected with the virus MMTV."
 /clone_lib="NCI CGAP_Lu29"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"

ORIGIN

Alignment Scores:

Pred. No.: 5,75e-69 Length: 3025
 Score: 661.50 Matches: 146
 Percent Similarity: 52.73% Conservative: 47
 Best Local Similarity: 39.89% Mismatches: 146
 Query Match: 35.47% Indels: 27
 DB: 11 Gaps: 8

US-09-466-778B-11 (1-353) x BC049247 (1-3025)

QY 2 ThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeuAsn 21
 Db 1612 ACCGTCGCAACACACACGCGCTGTGAATGCCAGTAGGCTACGTGGGTGACG 1671
 QY 22 Cys---GluProGluGlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis 40
 Db 1672 TGTCTAGAGAGAGCTTGACCCCTGTGGACAGATGCTTGGAGAGATCATCATCTTGCAC 1731
 QY 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60
 Db 1732 ACTGATGCTTTGTGCACTGACCTACATTTCCAGGAAAAACAGCTGTGTCTTCCACATC 1791
 QY 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80
 Db 1792 CAGGCCACACAGTGGCCCTTATGCTCTGACCTTCTCAGAGGCCACGAGCGTGTAGGGC 1851
 QY 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr**GlnLysAlaLysTyrHis 100
 Db 1852 CAGGAGCGCTCTGCTTCACTCCCTCAACTCTCTGCGGCCAACAGCTAGGTTTCAT 1911

JOURNAL
COMMENT

Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics / NIH
National Cancer Institute
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: gcapbs-i@mail.nih.gov
Tissue Procurement: Len Zon, Harvard
cDNA Library Preparation: Open Biosystems
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: L14M14828 row: b column: 03
High quality sequence stop: 751.
Location/Qualifiers

FEATURES
source

1. 1122
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:7053029"
/tissue_type="whole body"
/lab_host="DH10B"
/clone_lib="NIH ZGC 7"
/note="Vector: pExpress1; Site 1: NotI; Site 2: EcoRV;
Bulk tissue was collected from a whole adult individual
from the Tuebingen strain. 1st strand cDNA was primed with
a Not I - oligo(dT) primer, double-stranded cDNA was
cloned into the Not I and EcoRV sites of pExpress-1.
Library was size-selected for >1 kb fragments and
normalized. A non-normalized version of this library is
also available (NIH ZGC 10). Library was constructed by
Open Biosystems (Huntsville, AL)"

ORIGIN

Alignment Scores:
Pred. No.: 8.07e-67 Length: 1122
Score: 638.00 Matches: 132
Percent Similarity: 59.75% Conservative: 56
Best local similarity: 41.25% Mismatches: 117
Query Match: 34.21% Indels: 16
DB: 14 Gaps: 5

US-09-466-778B-11 (1-353) x CK028027 (1-1122)

QY	101	LeuCySerAlaGlyTrpLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer	120
Db	1912	GTCTGCTTCGCGGGCTGGTGGCCAAACGGCTCTGCTGCCACCTGTGTGTCACCCAGCG	1971
QY	121	GlnAenCysGlySerGlyValValGlyLeuValAspTyrGlyProArgProAsnLysSer	140
Db	1972	GCAGACTGTGTAATAATCGTGTAGGTAGTATGACCTTGGGGCCGTAAGAACTCTTCG	2031
QY	141	GluMetTrpAspValPheCysTyrArgMetLysAspValAenCysThr**LysValGly	160
Db	2032	GAGCTCTGGATGCTACTTATCCGCTGCAAGACGTGGCTTGCAGTGTGCGGCGCGC	2091
QY	161	TyrValGlyAspGlyPheSer---TyrSerGlyAenLeuLeuGlnValLeuMetSerPhe	179
Db	2092	TTCGTGGGTGACGGGATCAGCACGTGCAACGGGAACCTGCTTGTATGCTTGGCGCCACT	2151
QY	180	ProSerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGly	199
Db	2152	GCCAACTTCTCACCTTCTATGGGATGCTGTGGGCTATGCCAATGCCACCCAGAGGT	2211
QY	200	ArgAlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGln	219
Db	2212	CTTGATTTATGATTTCTTGAGGATGAGCTACCTACAGACACTCTTGTTCTCTGTC	2271
QY	220	AsnSerGlyLeuGlyGluAenGluThrLeuSerGlyArgAspIleGluHisLeuAla	239
Db	2272	AACAAAGGCTTTGTGGACAAACATGACGCTGAGTGCCGACATCTAGAACTCCACGCTTCT	2331
QY	240	AsnValSerMetPhePheTyrAsnAspLeuValAenGlyThrLeuGlnThrArgLeu	259
Db	2332	AATGCCACTTCTGAGTGTCAAT---GCCACCGGGGGACATGCTTCTCTGCCACTCA	2388
QY	260	GlySerLysLeuLeuThrAspArg-----GlnAspProLeuHisProThrGlu	276
Db	2389	GGTCTTAGCTCTTCATAGTACATGAGTGGCCCTGACACACTCTCTGTTCTCTGGCC	2448
QY	277	ThrArgCysValAspGlyArgAspThrLeuGluTrpAspIleCysAlaSerAenGly	296
Db	2449	CGGGGGGCGAGTTGTGGTTAGCCACGTCATCGTGTGGACATCATGCTTTCACGCGATC	2508
QY	297	ThrHisValIleSerArg**LeuLysAlaPro-----	307
Db	2509	ATTATGCTCTGCGCAGCCCTCTCATGCTCCCTCCACACTAGGCGAGCTGTGGATCT	2568
QY	308	---ProAlaProValThrLeu**HisThrGlyLeuGly**GlyLePhe*****Ile	326
Db	2569	GAGCTCCACCTGTGGCATT-----AGCTTGGGGGTGTGTAACCTTCTGGAACA	2619
QY	327	Ile-----LeuValThrGlyAlaValAlaLeuAlaTyrSerTyrPheArgIleAen	344
Db	2620	CTGCTGGGCTGTGTCGGGAGCTCTC-----TACCTGGCTGCGCGCA	2661
QY	345	ArgLysThrIleGlyPhe	350
Db	2662	GGCAAACTCCAGGTTC	2679

RESULT 6
CK028027
LOCUS
DEFINITION AGENCOURT_16620198 NIH_ZGC_7 Danio rerio linear EST 26-NOV-2003
5', mRNA sequence.
CK028027
CK028027.1 GI:38553951
VERSION
KEYWORDS
SOURCE
ORGANISM
Danio rerio (zebrafish)
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 1122)
REFERENCE
AUTHORS
TITLE
NTH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)


```

QY 142 MetTrpAspValPheCysTyrArgMetLysAspValAsnCysThr***LysValGlyTyr 161
Db 422 CTGTACAGTCGTACTGCTACAGATGAATGAGTGGCTTGTGAATGTGGAGCTGGATAT 481
QY 162 ValGlyAspGlyPheSerTyrSerGlyAsnLeuGlnValLeuMetSerPheProSer 181
Db 482 GTTGGAGCGGGCATTTTGTAAATGGAGACTTGGCTTGGTTGTGGCCAACTTCCAAAT 541
QY 182 LeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArgAla 201
Db 542 TTTTCTGTGTTTACACGAATTTACTGAAGTACGACGAGGATGGAGAGGACGCAAGACC 601
QY 202 PheLeuGlnHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsnSer 221
Db 602 CTGCTGAAGGCCCTGTCTGCAATCAATCAATTAACATGACCACTTTTATACCCCAAGAACAC 661
QY 222 GlyLeuGlyGluAsnGlnThrLeuSerGlyArgAspIleGluHisLeuAlaAsnVal 241
Db 662 GGATTTGGAGNCACAGACTTTATCTGTTGGAGATTTGCGATATCATATCTCAGGGTC 721
QY 242 ---SerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGly 260
Db 722 AACAAATATTCACTATTATGAGAACCTGAAGGCACACACTCAATCCCGTCTCGCTGGGC 781
QY 261 SerIleValLeuIleThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal 280
Db 782 TC-AAACTCATCAATCAACATCCCTCCATCCCA-----CGGTCGACACTGAGGAGCGTC 834
QY 281 AspGlyArgAspThrLeuGluTyrAspIle-----CysAla 292
Db 835 AACCTCAGA-----ACCTGGGATTAAGAGGTATTTTGGCTGAACCATTCGGCC 885
QY 293 SerAsnGlyIleThrHisValIleSerArg***LeuLysAlaProProAlaProValThr 312
Db 886 ATCAATGGACTGATCCAGCTTCATGATGGGCTTTCAAACTCCCTCGTGGTTCCTTC 945

```

```

RESULT 7
BX739778 722 bp mRNA linear EST 18-NOV-2003
LOCUS BX739778 XGC-tadpole silurana tropicalis cDNA clone TTPA048d24 5',
DEFINITION mRNA sequence.
ACCESSION BX739778
VERSION BX739778.1 GI:38412519
KEYWORDS EST.
SOURCE Silurana tropicalis (western clawed frog)

```

```

ORGANISM Silurana tropicalis (western clawed frog)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Silurana.
1 (bases 1 to 722)
Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
Sanger Xenopus tropicalis EST project 2001 (11_2003)
Unpublished (2003)
Contact: Croning MDR
Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS SEQUENCE ID: TTPA048d24.plkasp6
Sequencing primer: Sp6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Nigel Garrett.
cDNA was oligo dt primed from 5ug of poly A+ RNA from tadpole
embryos. EcoRI-NotI cut cDNA was then ligated into pCS107 with
EcoRI at the 5' end and NotI at the 3' end.
Vector: pCS107; Site 1: EcoRI; Site 2: NotI
Host: Escherichia coli DH10B.
Location/Qualifiers
1..722
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="TTPA048d24"

```

```

FEATURES
source

```

```

/dev stage="tadpole (stage 35-40)"
/lab_host="E. coli DH10B"
/clone_lib="XGC-tadpole"
/notes="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
was oligo dt primed from 5ug of poly A+ RNA from tadpole
embryos. EcoRI-NotI cut cDNA was then ligated into pCS107
with EcoRI at the 5' end and NotI at the 3' end"

```

ORIGIN

```

Alignment Scores:
Align. No.: 4,92e-63 Length: 722
Score: 604.50 Matches: 119
Percent Similarity: 71.22% Conservative: 27
Best Local Similarity: 58.05% Mismatches: 52
Query Match: 32.41% Indels: 7
DB: 13 Gaps: 2

US-09-466-778B-11 (1-353) x BX739778 (1-722)

QY 152 AspValAsnCysThr***LysValGlyTyrValGlyAspGlyPheSerTyrSerGlyAsn 171
Db 1 GATGTGTCTGATGATTTCGAAACCTCCCTTTGTGGAGATGGCTACACTTGAATGAAGAAC 60
QY 172 LeuLeuGlnValLeuMetSerPheProSerLeuThrAsnPheLeuThrGluValLeuAla 191
Db 61 CTGCTCAAGTCTTGACATCTTTTCCCTCTCTTCAAAATTCCTTACGGAAATTTTGGTG 120
QY 192 TyrSerAsnSerSerAlaArgGlyArgAlaPheLeuGluHisLeuThrAspLeuSerIle 211
Db 121 TACTCTAACACTTTCATGAAAGGAAAGAAATTTTCAACTACCTCACTCACTTATCAGTG 180
QY 212 ArgGlyThrLeuPheValProGlnAsnSerGlyLeuGlyGluAsnGlnThrLeuSerGly 231
Db 181 CAGCGACTCTGTTTGGCCCCAAGCAATGATGGACTTAATGAAATCAGACCCCTGCTGGA 240
QY 232 ArgAspIleGluHisLeuAlaAsnValSerMetPhePheTyrAsnAspLeuValAsn 251
Db 241 CGCGATATAGAAATATCATCTGGCAATGTCCAGCATGTTCTTCTTGGAGCACTGAGCAAT 300
QY 252 GlyThrThrLeuGlnThrArgLeuGlySerLysLeuLeuIleThrAspArgGlnAspPro 271
Db 301 GGAAACACTCTTCAACACCAAGATAGGCCACAGTATTATTAATTTCTTTGACAAATGACCCA 360
QY 272 LeuHisPro-----ThrGluThrArgCysValAspGlyArgAspThr 285
Db 361 GCCAGCAAAATAACCACTGATTCCATGACTCCCAACAGATATGTGGATGGAAACCTATT 420
QY 286 LeuGluTyrAspIleCysAlaSerAsnGlyIleThrHisValIleSerArg***LeuLys 305
Db 421 TTGCAGTGGGACATTATTCGTTTCCAAATGGAATCATTCACACAATTTGCTGAACCTCTAACT 480
QY 306 AlaProProAlaProValThrLeu***HisThrGlyLeuGly***GlyIlePhe***** 325
Db 481 GCGCTCCAGACCACTAGCTCTG---CATGCTGGCCATGGAGCAGGAATATTTCTTTGGC 537
QY 326 IleIleLeuValThrGlyValAlaLeuAlaTyrSerTyrPheArgIleAsnArg 345
Db 538 ATTGTCTGATCGTGGGCTCTGGCACTCGCAGTATATACTATTAAGAAGTTCAACAGA 597
QY 346 LysThrIleGlyPhe 350
Db 598 AAAGATTTCCAATTC 612

```

```

RESULT 8
BX732486
LOCUS BX732486.1 GI:15944036
DEFINITION mRNA sequence.
ACCESSION BX732486
VERSION BX732486.1 GI:15944036
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

```

```

BX732486 793 bp mRNA linear EST 04-OCT-2001
603082278F1 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5221410 5',
mRNA sequence.

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 793)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Inyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM11556 row: 1 column: 19

High quality sequence stop: 783.

FEATURES

Location/Qualifiers

1..793

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5221410"

/lab_host="DH10B"

/clone_lib="NIH_MGC_120"

/note="Organ: pooled pancreas and spleen; Vector:

pcmw-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA

source anonymous pool of spleen and pancreas from 28 yo

male. Library is oligo-dr primed and directionally cloned

(EcoRV site is destroyed upon cloning). Average insert

size 1.5 kb, insert size range 1-2.5 kb. Library is

normalized and enriched for full-length clones and was

constructed by C. Gruber (Invitrogen). Research Genetics

tracking code 025. Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.: 3,49e-61 Length: 793

Score: 590.00 Matches: 127

Percent Similarity: 63.68% Conservative: 1

Best Local Similarity: 63.18% Mismatches: 8

Query Match: 31.64% Indels: 66

DB: 12 Gaps: 1

US-09-466-778b-11 (1-353) x BI832486 (1-793)

QY 91 LeuSerTyr***GlnLysAlaLysTyrHisLeuCysSerAlaGlyTrpLeuGluThrGly 110

DB 193 CTGTCTGTATCCTCGAGGCAAGTACCACCTGTCTCAGCAGGCTGGCTGGAGACCGGG 252

QY 111 ArgValAlaTyrProThrAlaPheAlaSerGlnAsnCysGlySerGlyValValGlyIle 130

DB 253 CGGGTTGGCTACCCACAGCCTTCGCTCCAGAACTGTGGCTCTGTGGTGGGATA 312

QY 131 ValAspTyrGlyProArgProAsnLysSerGluMetTrpAspValPheCysTyrArgMet 150

DB 313 GTGGACTATGGACCTAGACCCAAAGAGTGAATGTGGATGCTCTTCGTATCGGATG 372

QY 151 Lys-AspValAsnCysThr***LysValGlyTyrValGlyAspGlyPheSerTyrSerG1 170

DB 373 AACAGATGTGAACCTGCACCTGCAAGTGGGCTATGTGGGAGATGCTTCTCATGCAGTGG 432

QY 170 YAsnLeuLeuGlnValLeuMetSerPheProSerLeuThrAsnPheLeuThr----- 187

DB 433 GAACCTGTGTGAGTCTCTGATGCTTCCCTCACTCAAACTTCTGACGGATTCACT 492

QY 187 ----- 187

DB 493 GAGGAAGAAAGTCTCTCAAACTCTCCACGGCGGAGGAGACAGACAGCAAGGAGC 552

QY 187 ----- 187

DB 553 AGCAGCTGAGTTGGCTCACCATGATATTCACAGATGGGTGATGCCCTCGGACCTTC 612

QY 187 ----- 187

DB 613 TTCAGCACCCAGGTCCCTGCACACACAGGCTGGCCCTCTGTATGACCTCTTCTTCTCC 672

QY 188 ----GluValLeuAlaTyrSerAsnSerAlaArgGlyArgAlaPheLeuGluHisLe 206

DB 673 CCAGGAAGTGTGGCCCTATTTCCACAGCTCAGCTCGAGGCGTGCAATTTTAGACCACT 732

QY 206 uThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsnSerGlyLeuGlyGluAs 226

DB 733 GACTGACCTGTCCATCCGCGGACCTC-TTTGTGCCACAGACAGTGGCTGGGGAGAA 791

QY 226 n 226

DB 792 T 792

RESULT 9

BC019712 2368 bp mRNA linear HTC 20-SEP-2002

LOCUS Mus musculus, Similar to stabilin 1, clone IMAGE:3663004, mRNA.

DEFINITION BC019712

ACCESSION BC019712.1 GI:22507492

VERSION HTCC

KEYWORDS Mus musculus (house mouse)

SOURCE Mus musculus

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 2368)

AUTHORS Strausberg, R.

TITLE Direct Submission

JOURNAL Submitted (19-DEC-2001) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk

Email: cgapbs-remail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Baylor College of Medicine Human Genome

Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: amg@bcm.tmc.edu

Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,

Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,

Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK Plate: 40 Row: d Column: 23

This clone has the following problem: frame shifted.

Location/Qualifiers

1..2368

/organism="Mus musculus"

/mol_type="mRNA"

/strain="CZECH II"

/db_xref="taxon:10090"

/clone="IMAGE:3663004"

/tissue_type="Mammary tumor metastasized to lung.

MMTV-LTR/Wnt1 model. Expression driven by an MMTV-LTR

enhancer."

/clone_lib="NCI CGAP_Lu30"

/lab_host="DH10B"

/note="Vector: pcMV-SPORT6"

ORIGIN

Alignment Scores:

Pred. No.: 1.37e-58 Length: 2368

Score: 575.50 Matches: 147

Percent Similarity: 43.72% Conservative: 48

```

Best Local Similarity: 32.96% Mismatches: 144
Query Match: 30.86% Indels: 109
DB: 11 Gaps: 11
US-09-466-778B-11 (1-353) x BC019712 (1-2368)

QY 2 ThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeuAsn 21
DB 725 ACCGGTCCGAAACACACGCGCGTGTGAATGCCACATAGCCATAGCTACGTGGGTGATGGGCTCAG 784
QY 22 Cys---GluProGluGlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis 40
DB 785 TGCTAGAGGAGCTTGAACCCCTGTGGACAGATGCTTGGAGGATCATCACCTTGGCCAC 844
QY 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60
DB 845 ACTGATGCTTTGTGCACTGACCTACATTTCCAGGAAAAACAGGCTGGTGTCTTCCACATC 904
QY 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80
DB 905 CAGGCCACCAAGTGGCCCTTATGGTCTGACCTTCACAGAGCCACGGAAGCGTGTGAGGGC 964
QY 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis 100
DB 965 CAGGAGCCGTCCTTGTCTCACTCCCTCACTCTCTGCGGCCCAACAGCTAGGTTTCAT 1024
QY 101 LeuCysSerAlaGlyTyrLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120
DB 1025 GTCTGCTTGGGGCTGGTGGCCACAGCGCTCTGCTGCCACCTGTGTGTCACCCACGCG 1084
QY 121 GlnAsnCysGlySerGlyValValGlyLysValAspTyrGlyProArgProAsnLysSer 140
DB 1085 GCAGCTGTGGTAATATCGTGTAGGTAGTGTAGTTCCTTGGGTCGCGTAAGAACCTCTCG 1144
QY 141 GluMetTyrAspValPheCysTyrArg-----Met 150
DB 1145 GAGCTCTGGGATGCTACTGTACCGGTGCAAGGTAGCTCCCTTCGTCGATCTCTGCC 1204
QY 1205 TACTCTCGGCCCTCGAGCCCTTAATAAGCCTCTGTGAGCCACTGACCCGCTTTCCTG 1264
QY 151 LysAspValLeuAsnCysThr***LysValGlyTyrValGlyAspGlyPheSer---TyrSer 169
DB 1265 CA-GACGTGGCTTGCAGTGTCCGCGCGCTCTGCTGGGTGACGGATCAGCAGCTGCAAC 1323
QY 170 GlyAsnLeuLeuGlnValLeuMetSerPheProSerLeuThrAsnPheLeuThrGluVal 189
DB 1324 GGGAACTGCTTGATGCTTGGCGCCACTGCGCACTTCTCCACTTCTATGGGATGCTG 1383
QY 190 LeuAlaTyrSerAsnSerSerAlaArgGlyArgAlaPheLeuGluHisLeuThrAspLeu 209
DB 1384 CTGGCTATGCAATGCGCCACGAGAGGTCTTGAATTTATGATTTCTCTGGAGGATGAG 1443
QY 210 SerIleArgGlyThrLeuPheValProGlnAsnSerGlyLeuGlyGluAsnGlu----- 227
DB 1444 CTCACCTACAAGACACTCTTCGTTCTCTCAAAAGGCTTTGGGACAAATGGTAACG 1503
QY 227 ----- 227
DB 1504 GCAGGGCTTGGCGAGTTGCTCTCTGCACTCGGCCCAACCCACAGGCTCTCTCTCTGC 1563
QY 228 -----ThrLeuSerGlyArgAspIleGluHisLeuAlaAsnValSerMetPhePhe 245
DB 1564 TCATAGACGCTGAGTGGCCAGATCTAGAACTCCACGCTTCTAATGCCACCTTCTGAGT 1623
QY 246 TyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGlySerLysLeuLeuIle 265
DB 1624 GTCAAT---GCCAGCGGGGACATATGCTTCTCTGCCCACTCAGGTCTTAGCCTCTTCA 1680
QY 266 ThrAspArgGlnAspProLeuHis----- 273
DB 1681 AGTGACATGGGCCCTGACAA-CACTTCTCTGGTTCCTCTGTGCGTCTGCCCTCACTGA 1739

```

```

QY 274 ---ProThrGluThrArg----- 278
DB 1740 AACCTTACTGACTCAGCCCTTGTCTACCGGCTGGGCTGACTGAGTTTCTCCATAGGCC 1799
QY 279 -----CysValAspGlyArgAspThrLeuGluTyrAspIleCysAlaSerAsnGlyIle 296
DB 1800 CCGGGGCGAGCTGTGGTTAGCCACGTCATCGTGGGACATCATGGCTTTCAACGGCATC 1859
QY 297 ThrHisValIleSerArg***LeuLysAlaPro----- 307
DB 1860 ATTCACTCTGCGCCACCCCTGCTCATGCTCCCGACTAGGCGAGTGTGGATCT 1919
QY 308 ---ProAlaProValThrLeu***HisThrGlyLeuGly***GlyIlePhe*****Ile 326
DB 1920 GAGCTCCACCTGTGGCATTA-----AGCTTGGGGCTTGTGTAACCTTCTGGAACA 1970
QY 327 Ile-----LeuValThrGlyAlaValAlaLeuAlaLysTyrSerTyrPheArgIleAsn 344
DB 1971 CTGCTGGGGCTGTGGCGGAGCTCTC-----TACCTGGCTGGCCGA 2012
QY 345 ArgLysThrIleGlyPhe 350
DB 2013 GGCACCACTCCAGGTTTC 2030
RESULT 10
T47504
LOCUS
DEFINITION
T47504
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT
FEATURES
source

```

T47504 459 bp mRNA linear EST 01-FEB-1995
 yb14f01.r1 Stratagene placenta (#937225) Homo sapiens cDNA clone
 IMAGE:71161 5', similar to similar to SP:A41735 A41735
 HYALURONATE-BINDING PROTEIN TSG-6 PRECURSOR, mRNA sequence.

T47504 GI:649484
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 459)
 Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,
 Chisoe, S., Dietrich, N., DuBuque, T., Favello, A., Gish, W.,
 Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,
 Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,
 Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,
 Trevaskis, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R.
 and Marra, M.
 Generation and analysis of 280,000 human expressed sequence tags
 Genome Res. 6 (9), 807-828 (1996)
 97044478
 8889349
 Other ESTs: yb14f01.s1
 Contact: Willson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.wustl.edu
 Insert Size: 1402
 High quality sequence stops: 308 Source: IMAGE Consortium, LLNL This
 clone is available royalty-free through LLNL; contact the IMAGE
 Consortium (info@image.llnl.gov) for further information.
 Insert length: 1402 Std Error: 0.00
 Seq primer: M13RP1
 High quality sequence stop: 308.
 Location/Qualifiers
 1..459
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="GDB:492058"
 /db_xref="taxon:9606"
 /clone="IMAGE:71161"
 /sex="male"
 /lab_host="SOLR cells (kanamycin resistant)"

```
/clone_lib="Stratagene placenta (#937225)"
/notes="Organ: placenta; Vector: pBluescript SK-; Site 1:
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. Caucasian. Average insert size: 1.2 kb; Uni-ZAP
XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGGAG 3' -3'
adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"
```

ORIGIN

Alignment Scores:
Pred. No.: 1.34e-57 Length: 459
Score: 557.50 Matches: 128
Percent Similarity: 85.06% Conservative: 3
Best Local Similarity: 83.12% Mismatches: 16
Query Match: 29.89% Indels: 9
DB: 14 Gaps: 1

US-09-466-778B-11 (1-353) x T47504 (1-459)

Qy 88 TyrAsnGlnLeuSerTyr**GlnIysAlaLysTyrHisLeuCysSerAlaGlyTrieu 107
Db 2 TACAAACAGCTCTCTCTATGCCAGAGGC-AAGTACCACCTGTCTCAGCAGGCTGGCTG 60
Qy 108 GlnThrGlyArgValAlaTyrProThrAlaPheAlaSerGlnAsnCysGlySerGlyVal 127
Db 61 GAGACCGGGGGTTCCTACCCACAGCTTCGCTCCAGAACTGGCTGGCTGGTGTG 120
Qy 128 ValGlyIleValAspTyrGlyProArgProAsnLysSerGluMetTyrAspValPheCys 147
Db 121 GTTGGGATGATGGACTATGGACCTAGACCCACAGAGTGAATGTGGGATGTCTTCCTG 180
Qy 148 TyrArgMetIysAspValAsnCysThr**IysValGlyTyrValGlyAspGlyPheSer 167
Db 181 TATCGGATGAAGATGTGAACCTGACCTGCACCTG-CAAGTGGCTATGTGGGAGATGGCTTCTCA 239
Qy 168 TyrSer-GlyAsnLeuLeuGlnValLeuMetSerPheProSerLeuThrAsnPheLeuTh 187
Db 240 TGCAGTGGGAACTGCTGCAGTCTCTGATGCTCTCTCCCTCACTCACAAATCTTCGTAC 299
Qy 197 rGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArgAlaPheLeu--GluHisLeu 206
Db 300 GGAAGTGTGGCTATTCCACAGCTCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCC 359
Qy 207 ThrAspLeu-SerIleArgGlyThrLeuPhe--ValProGlnAsnSerGlyLeuGlyCln 225
Db 360 ATTGACCTGTTCATCTCCCGGGCAGCTTTTGTGTGCCACAGAACAGTNGGCTTGGGGG 419
Qy 226 Asn---GluThrLeuSerGlyArgAspIle 234
Db 420 AGATTGAGACNTTTTGGGGGGGACATC 449

RESULT 11
AK028441
LOCUS
DEFINITION
Mus musculus 0 day neonate skin cDNA, RIKEN full-length enriched library, clone:4631403B18 product:hypothetical Beta-Ig-H3/Fasciclin domain containing protein, full insert sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
AK028441
AK028441.1 GI:26080850
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
10349636
2
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
10349636
2
Carninci, P., Shibata, Y., Hayateu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

TITLE

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3454)

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Horii, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akai, S., Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/

FEATURES
source

1. 3454
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM:DB:4631403B18"
/db_xref="MGI:2330349"
/db_xref="taxon:10090"
/clone="4631403B18"
/tissue_type="skin"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="0 day neonate"
1. 3454

misc_feature

/note="hypothetical Beta-Ig-H3/Fasciclin domain containing protein (InterPro|IPR000782, evidence: InterPro)"

ORIGIN

US-09-466-778B-11 (1-353) x BX446538 (1-1201)

Qv 18 AspGlvLeuAsnCys---GluProGluGlnLeuproileAspArgCysLeuGlnAspAsn 36

Db 29 GATGACTGCAGTGTCTGGAGAGTGGGAACACACTGTGGACCGTGTCTGGCCAGCCA 88
 Qy 37 GlyGlnCysHisAlaAspAlaLysCysValAspLeuHisPheGln----- 51
 Db 89 CGCCCTCCCACTCAGATGCATGTGCTGACCTGACCTCCAGGGTGTGTCCTCCCTGC 148
 Qy 51 ----- 51
 Db 149 CCATCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 208
 Qy 52 -----AspThrValValGlyValPheHisLeuArgSerProLeuGlyGlnTyrLysLeu 69
 Db 209 TCTGCAGAAACGGCTGGCTGCTTTCCTCCCTCCAGCCAGCCAGCGGCTTATGCTGTG 268
 Qy 70 ThrPheAspLeuAlaArgGluAlaCysAlaAsnGluAlaThrMetAlaThrTyrAsn 89
 Db 269 AACTTTTCGGAGCTGAGCGCATGGGA-ASACAGGAGCGCTGCTTCTTCATTCCT 327
 Qy 90 GlnLeuSerTyr***GlnLysAlaLysTyrHisLeuCysSerAlaGlyTrpLeuGluThr 109
 Db 328 CAGCTCTCTGCTGCCAGCAGCTGGCTTCCACCTGTGCTCATGGCTGGCTGGCCAA 387
 Qy 110 GlyArgValAlaTyrProThrAlaPheAlaSerGlnAsnGlySerGlyValValGly 129
 Db 388 GGTCTCCACTGCCCTGCTGCTTTCCTGTGGCGGACTGTGGCAATGGTGGGTGGGC 447
 Qy 130 IleValAspTyrGlyProArgProLysSerGluMetTrpAspValPheCysTyrArg 149
 Db 448 GTAGTCAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 507
 Qy 150 MetLysAspValAsnGlyThr***LysValGlyTyrValGlyAspGlyPheSer---Tyr 168
 Db 508 GTGCAAGATGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 567
 Qy 169 SerGlyAsnLeuLeuGlnValLeuMetSerPheProSerLeuThrAsnPheLeuThrGlu 188
 Db 568 AATGGGAAGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 627
 Qy 189 ValLeuAlaTyrSerAsnSerSerAlaArgGlyArgAlaPheLeuGluHisLeuThrAsp 208
 Db 628 CTATTGGGTATGCCAATGCCAGCCAGCGGGTCTGCACTTCTGCACTTCTGCACTTCTGCACT 687
 Qy 209 LeuSerIleArgGlyThrLeuPheValProGlnAsnSerGlyLeuGluAsnGluThr 228
 Db 688 GAGCTCACATTAAGACATCTTCGCTCCCTGCTCAATGAAGCTTGTGGACAAATGACG 747
 Qy 229 LeuSerGlyArgAspIleGluHisLeuAlaAsnValSerMetPheTyrAsnAsp 248
 Db 748 CTGAGTGGCCAGACTTGGAGCTGATGCTCCAGCCAGCCCTCCTAAGTGCCCAAC--- 804
 Qy 249 LeuValAsnGlyThrThrLeuGlnThrArgLeuGlySerLysLeuLeuLeuThrAspArg 268
 Db 805 GCACGACGAGGAGGAGTGTCTTCGGCCCACTCAGGCTCAGGCTCATCATCATGACGCA 864
 Qy 269 GlnAspProLeuHisProThrGluThrArgCysValAspGlyArgAspThrLeuGluTrp 288
 Db 865 -----GGCCCTGACACAGTCTCTGG 885
 RESULT 13
 CB268774
 LOCUS
 DEFINITION 572 bp mRNA linear EST 20-FEB-2003
 CDNA 5' mRNA sequence.
 ACCESSION CB268774
 VERSION CB268774.1 GI:28443359
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 572)
 REFERENCE Yang,R.-Z., Shuldiner,A. and Gong,D.-W.
 AUTHORS
 TITLE EST analysis of human adipose gene expression

JOURNAL COMMENT
 Unpublished (2002)
 Contact: Gong Da-Wei
 Division of Endocrinology, Diabetes and Nutrition
 University of Maryland
 660 Redwood St, HH497, Baltimore, MD 21201, USA
 Tel: 410 706 1672
 Fax: 410 706 1622
 Email: dgong@medicine.umaryland.edu
 PCR Primers
 FORWARD: CTGGGAAGCGCCCATTTGTGTGGT
 BACKWARD: AATACGACTCATTAGGCGCAATTGG
 Seq primer: GTTGTACCGGGAATTC.
 Location/Qualifiers
 1..572
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /sex="Male and Female"
 /tissue_type="Adipose"
 /clone_lib="Human Fat Cell 5'-Stretch Plus cDNA Library"
 /note="Vector: lambdaTriplex"

FEATURES
 source
 1..572

ALIGNMENT Scores:
 Pred. No.: 1.04e-45 Length: 572
 Score: 462.00 Matches: 86
 Percent Similarity: 62.57% Conservative: 31
 Best Local Similarity: 45.99% Mismatches: 68
 Query Match: 24.77% Indels: 2
 DB: 14 Gaps: 2

ORIGIN
 US-09-466-778B-11 (1-353) x CB268774 (1-572)
 Qy 15 TyrValGlyAspGlyLeuAsnGly---GluProGluGlnLeuProIleAspArgCysLeu 33
 Db 6 TACGTAGCGATGAGTGTCTGAGGAGTGGGAACCCCTGTGGACCGCTGCTG 65
 Qy 34 GlnAspAenGlyGlnCysHisAlaAspAlaLysCysValAspLeuHisPheGlnAspThr 53
 Db 66 GGCAGCCAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 125
 Qy 54 ThrValGlyValPheHisLeuArgSerProLeuGlyGlnTyrLysLeuThrPheAspLys 73
 Db 126 CGGCTGCGGCTTCCACCTCCAGCCAGCCAGCCCTTATGCTGCTGCTGCTGCTGCTGCT 185
 Qy 74 AlaArgGluAlaCysAlaAsnGluAlaThrMetAlaThrTyrAsnGlnLeuSerTyr 93
 Db 186 GCTGAGCGCGATGCGAAGACAGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 245
 Qy 94 ***GlnLysAlaLysTyrHisLeuCysSerAlaGlyTrpLeuGluThrGlyArgValAla 113
 Db 246 GCCCAGCAGCTGGGCTTCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 305
 Qy 114 TyrProThrAlaPheAlaSerGlnAsnGlySerGlyValValGlyIleValAspTyr 133
 Db 306 CACCTGCGGCTTCCCGTGGCGACTGTGCAATGTGCGGTGGGCGGTAGTCAGCCTG 365
 Qy 134 GlyProArgProAsnLysSerGluMetTrpAspValPheCysTyrArgMetLysAspVal 153
 Db 366 GGTGCGCCGCAAGACCTCTCAGAACGCTGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 425
 Qy 154 AsnGlyThr***LysValGlyTyrValGlyAspGlyPheSer---TyrSerGlyAsnLeu 172
 Db 426 GCCTGCCATGCCGAATGGCTTGTGGTACGGGATCAGCAGCTGCAATGGGAGCTG 485
 Qy 173 LeuGlnValLeuMetSerPheProSerLeuThrAsnPheLeuThrGluValLeuAlaTyr 192
 Db 486 CTGGATGTGCTGGCTGCCACTGCCACTCTCCACCTTCTATGGATGCTATGGCTAT 545
 Qy 193 SerAsnSerSerAlaArgGly 199
 Db 546 GCCAATGCCCAACACGGGT 566

```

Db      241  ATGACTCCACAGATATGCGATGGAAAACCTATTTTGCAGTGGGACATTATGCTCTCC 300
QY      294  AenGlyIleThrHisValIleSerArg***LeuLysAlaProProAlaProValThrLeu 313
Db      301  AATGGAATCATTCACACAATTTGCTGAACCTCTAACTCGCGCTCCAGAGCCATTAGCTCTG 360
QY      314  ***HisThrGlyLeuGly***GlyIlePhe*****IleIleLeuValThrGlyAlaVal 333
Db      361  ---CATGCTGGCATTGAGCAGCAGGAAATATCTTTGGCAATTTGTTGATCTCGGCTCGT 417
QY      334  AlaLeuAlaAlaTyrSerTyrPheArgIleAsnArgLysThrIleGlyPhe 350
Db      418  GCACCTGCGAGTATACTATTATAGAAGATTCAACAGAAAAGATTTCCAATTC 468

RESULT 15
CA342417      719 bp      mRNA      linear      EST 05-NOV-2002
LOCUS       672284 NCCOWA lrt Oncorhynchus mykiss cdna clone lrt32P18_D_H09 5
DEFINITION  mRNA sequence.
ACCESSION   CA342417
VERSION     CA342417.1 GI:24587588
KEYWORDS    EST.
SOURCE      Oncorhynchus mykiss (rainbow trout)
ORGANISM    Oncorhynchus mykiss
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE   1 (bases 1 to 719)
AUTHORS     Rexroad,C.E. and Keele,J.W.
TITLE       Sequence analysis of a rainbow trout normalized cdna library
JOURNAL     Unpublished (2002)
COMMENT     Contact: Rexroad CE
            USDA, ARS, National Center for Cool and Cold Water Aquaculture
            11876 Leetown Road, Kearneysville, WV 25430, USA
            Tel: 304 724 8340 x2129
            Fax: 304 725 0351
            Email: crexroad@nccwa.ars.usda.gov
            Single pass sequencing. Bases called with phred v0.020425.c and
            trimmed with the aid of the trim_alt option. Vector identified by
            cross match v0.990329.
FEATURES             Location/Qualifiers
     source           1..719
                     /organism="Oncorhynchus mykiss"
                     /mol_type="mRNA"
                     /db_xref="taxon:8022"
                     /clone="lrt32P18_D_H09"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /clone_lib="NCCOWA lrt"
                     /note="vector: pcmw SPORT6; Site 1: NotI; Site 2: SalI;
                     Library made from pooled tissue from brain, gill, liver,
                     spleen, muscle, and kidney."

ORIGIN
Alignment Scores:
Pred. No.:      5.84e-43      Length:      719
Score:          441.00      Matches:      96
Percent Similarity: 56.90%      Conservative: 40
Best Local Similarity: 40.17%      Mismatches: 97
Query Match:     23.65%      Indels:      6
DB:              14          Gaps:      2

US-09-466-778B-11 (1-353) x CA342417 (1-719)

QY      67  TyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsnGluAlaAlaThrMetAla 86
Db      2   TACCAGCTGAACATACACCGATGCTGAGCTAGCCTGTCAAGCTGAGGGGGCTACGCTGCC 61
QY      87  ThrTyrAsnGlnLeuSerTyr**GlnLysAlaLysTyrHisLeuCysSerAlaGlyTrp 106
Db      62  ACCTTCTCTCAGATGGGTGATGCACACAGCTGGGAATGCACCGATGTGTGTGTGGATGG 121

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 1, 2004, 11:22:04 ; Search time 1504 Seconds

(without alignments)
877.740 Million cell updates/sec

Title: US-09-466-778B-11

Perfect score: 1865

Sequence: 1 MTGPGHKCEKSHYVDGL.....ALAAYSYFRINRKTIGPHF 353

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2465228 seqs, 1869859620 residues

Total number of hits satisfying chosen parameters: 4930456

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09466778/runat_01042004_095644_10700/app.query.fasta_1.519
-DB=Published Applications NA -CFMT=fastap -SUFITX=p2n.rnpb -MINMATCH=0.1
-LOOPEXT=0 -UNIT8-bits -START=1 -END=1 -MATRIX=bls6m62
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pcet -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODEB=LOCAL -OUTFMT=ptc -NORM=ext -HAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09466778 @CEN 1 1 396 @runat_01042004_095644_10700
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

- 1: /cgn2_6/ptodata/2/pubnpa/US07_PUBCOMB.seq*
- 2: /cgn2_6/ptodata/2/pubnpa/PCT_NEW_PUB.seq*
- 3: /cgn2_6/ptodata/2/pubnpa/US06_NEW_PUB.seq*
- 4: /cgn2_6/ptodata/2/pubnpa/US06_PUBCOMB.seq*
- 5: /cgn2_6/ptodata/2/pubnpa/US07_NEW_PUB.seq*
- 6: /cgn2_6/ptodata/2/pubnpa/PCTUS_PUBCOMB.seq*
- 7: /cgn2_6/ptodata/2/pubnpa/US08_NEW_PUB.seq*
- 8: /cgn2_6/ptodata/2/pubnpa/US08_PUBCOMB.seq*
- 9: /cgn2_6/ptodata/2/pubnpa/US09A_PUBCOMB.seq*
- 10: /cgn2_6/ptodata/2/pubnpa/US09B_PUBCOMB.seq*
- 11: /cgn2_6/ptodata/2/pubnpa/US09C_PUBCOMB.seq*
- 12: /cgn2_6/ptodata/2/pubnpa/US09C_NEW_PUB.seq*
- 13: /cgn2_6/ptodata/2/pubnpa/US10A_PUBCOMB.seq*
- 14: /cgn2_6/ptodata/2/pubnpa/US10B_PUBCOMB.seq*
- 15: /cgn2_6/ptodata/2/pubnpa/US10C_PUBCOMB.seq*
- 16: /cgn2_6/ptodata/2/pubnpa/US10C_NEW_PUB.seq*
- 17: /cgn2_6/ptodata/2/pubnpa/US60_NEW_PUB.seq*
- 18: /cgn2_6/ptodata/2/pubnpa/US60_PUBCOMB.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
------------	-------	-------	--------	-------	-------------

1	1772	95.0	4576	9	US-09-842-930A-24	Sequence 24, Appl
2	1772	95.0	4642	14	US-10-133-172-3	Sequence 3, Appl
3	1772	95.0	4962	14	US-10-133-172-19	Sequence 19, Appl
4	1769	94.9	3260	15	US-10-028-248A-210	Sequence 210, Appl
5	1769	94.9	3260	15	US-10-107-782-210	Sequence 210, Appl
6	1558	83.5	8495	15	US-10-028-248A-3	Sequence 3, Appl
7	1558	83.5	8495	15	US-10-107-782-3	Sequence 3, Appl
8	1554	83.3	8444	15	US-10-028-248A-1	Sequence 1, Appl
9	1554	83.3	8444	15	US-10-107-782-1	Sequence 1, Appl
10	1514.5	81.2	4706	9	US-09-842-930A-1	Sequence 1, Appl
11	1514.5	81.2	4706	14	US-10-133-172-1	Sequence 1, Appl
12	661.5	35.5	3681	15	US-10-264-237-1320	Sequence 1320, Ap
13	645.5	34.6	2753	10	US-09-774-639-12	Sequence 12, Appl
14	645.5	34.6	2753	10	US-09-969-730-13	Sequence 13, Appl
15	645.5	34.6	2753	16	US-10-621-363-13	Sequence 13, Appl
16	642	34.4	2160	15	US-10-104-047-1824	Sequence 1824, Ap
17	632.5	33.9	1482	15	US-10-305-720-230	Sequence 230, Appl
18	581	31.2	2863	15	US-10-104-047-610	Sequence 610, Appl
19	363	19.5	474	10	US-09-918-995-32777	Sequence 32777, A
20	267.5	14.3	1144	13	US-10-044-090-446	Sequence 446, Appl
21	267.5	14.3	1144	14	US-10-247-671-124	Sequence 124, Appl
22	267.5	14.3	1422	15	US-10-295-027-291	Sequence 291, Appl
23	267.5	14.3	1430	15	US-10-295-027-1020	Sequence 1020, Ap
24	267.5	14.3	1728	13	US-10-044-090-445	Sequence 445, Appl
25	266.5	14.3	1414	9	US-09-799-118-1	Sequence 1, Appl
26	262	14.0	466	12	US-10-085-783A-58386	Sequence 58386, A
27	262	14.0	466	15	US-10-242-535A-58386	Sequence 58386, A
28	256.5	13.8	1411	9	US-09-765-231A-44	Sequence 44, Appl
29	248	13.3	351	14	US-10-060-036-880	Sequence 880, Appl
30	243.5	13.1	1414	15	US-10-133-937-18	Sequence 18, Appl
31	243.5	13.1	1414	15	US-10-295-027-293	Sequence 293, Appl
32	243.5	13.1	1414	15	US-10-159-563-18	Sequence 18, Appl
33	231	12.4	396	10	US-09-918-995-17767	Sequence 17767, A
34	213	11.4	718	14	US-10-023-896-38	Sequence 38, Appl
35	202.5	10.9	1365	13	US-10-098-841-328	Sequence 328, Appl
36	200.5	10.8	1985	9	US-09-909-320-212	Sequence 212, Appl
37	200.5	10.8	1985	9	US-09-909-088B-212	Sequence 212, Appl
38	200.5	10.8	1985	9	US-09-905-291A-212	Sequence 212, Appl
39	200.5	10.8	1985	9	US-09-902-853-212	Sequence 212, Appl
40	200.5	10.8	1985	9	US-09-907-824-212	Sequence 212, Appl
41	200.5	10.8	1985	9	US-09-907-841-212	Sequence 212, Appl
42	200.5	10.8	1985	10	US-09-904-011-212	Sequence 212, Appl
43	200.5	10.8	1985	10	US-09-906-742-212	Sequence 212, Appl
44	200.5	10.8	1985	10	US-09-906-838-212	Sequence 212, Appl
45	200.5	10.8	1985	10	US-09-907-613-212	Sequence 212, Appl

ALIGNMENTS

RESULT 1
US-09-842-930A-24
; Sequence 24, Application US/09842930A
; Publication No. US20020197681A1
; GENERAL INFORMATION:
; APPLICANT: Weigel, Paul
; TITLE OF INVENTION: Identification of Hyaluronan Receptor for Endocytosis
; FILE REFERENCE: 5820.603
; CURRENT APPLICATION NUMBER: US/09/842.930A
; CURRENT FILING DATE: 2001-04-22
; PRIOR APPLICATION NUMBER: 60/245,320
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: 60/199,538
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ IDS NOS: 56
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24
; LENGTH: 4576
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-842-930A-24

Alignment Scores:	3.72e-231	Length:	4576
-------------------	-----------	---------	------

Score: 1772.00 Matches: 334
Percent Similarity: 95.18% Conservative: 2
Best Local Similarity: 94.62% Mismatches: 17
Query Match: 95.01% Indels: 0
DB: 9 Gaps: 0

US-09-466-778B-11 (1-353) x US-09-842-930A-24 (1-4576)

Qy 1 MetThrGlyProGlyLysHisCysGluCysLysSerHisTyrValGlyAspGlyLeu 20
Db 2956 ATGACAGCCCGGGCAACCAAGTGTGAGTGAAGTCACTATGTCGGAGATGGGCTG 3015

Qy 21 AsnCysGluProGluLeuProLeuAspArgCysLeuGlnAspAsnGlyGlnCysHis 40
Db 3016 AACTGTGACCGGAGCAGCTGCCATTCACCGCTGCTTACAGGACAATGGCAGTGCAT 3075

Qy 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60
Db 3076 GCAGAGCCCAATGTGTCGACCTCCACTCCAGGATACACTGTGGGGTGTTCATCTA 3135

Qy 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80
Db 3136 CGCTCCCACTGGGCGCAGTATAAGCTGACCTTGACAAAGCCAGAGAGCCCTGTGCCAAC 3195

Qy 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis 100
Db 3196 GAAGCTGCGACCATGGCAACCTACACCAAGCTCTCTATGCCAGAGGCCAAGTACCCAC 3255

Qy 101 LeuCysSerAlaGlyTyrLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120
Db 3256 CTGTGCTCAGCAGCTGCTGGAGCCGGGGTGTGCTACCCACAGCCCTTGCCTCC 3315

Qy 121 GlnAsnCysGlySerGlyValValGlyLysValAspTyrGlyProArgProAsnLysSer 140
Db 3316 CAGAACTGTGCTGTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 3375

Qy 141 GluMetTyrAspValPheCysTyrArgMetLysAspValAsnCysThr***LysValGly 160
Db 3376 GAAATGTGGAGTGTCTCTGCTATCGGATGAAGATGAAGTGAAGTGAAGTGAAGTGAAG 3435

Qy 161 TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuLeuGlnValLeuMetSerPhePro 180
Db 3436 TATGTGGAGATGCTTCTCATGAGTGGGAACTGTGTCAGTCTCTGCTGCTCTCTCTCC 3495

Qy 181 SerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArg 200
Db 3496 TCACCTCAAACTTCTCTGACGGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3555

Qy 201 AlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsn 220
Db 3556 GCATTTCTAGAACACCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3615

Qy 221 SerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisHisLeuAlaAsn 240
Db 3616 AGTGGCTGGGGAGATGACACTTGTCTGGCGGGACATCGAGCACCACTCGCCCAAT 3675

Qy 241 ValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGly 260
Db 3676 GTCAGCATGTTTCTACAAATGACCTGTGCAATGGCACCACCCCTGCAAAACGAGGCTGGGA 3735

Qy 261 SerLysLeuLeuIleThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal 280
Db 3736 AGCAAGCTGCTCATCTGCGAGCCAGGACCCACTCCAAACCGAGGACCAAGGTTGTT 3795

Qy 281 AspGlyArgAspThrLeuLeuTriPAspIleCysAlaSerAsnGlyIleThrHisValIle 300
Db 3796 GATGGAAGAGCATTCTTCAGTGGGACATCTTGGCTCCATGGGATCATTCATGTCATT 3855

Qy 301 SerArg***LeuLysAlaProProAlaProValThrLeu***HisThrGlyLeuGly*** 320
Db 3856 TCCAGGCTTTAAAGAGCACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3915

Qy 321 GlyIlePhe****IleIleLeuValThrGlyAlaValAlaLeuAlaAlaTyrSerTyr 340

Db 3916 GGGATCTTCTTGGCATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3975

Qy 341 PheArgIleAsnArgLysThrIleGlyPhe***HisPhe 353
Db 3976 TTTCGGATAAACCGGAGAACAAATCGGCTTCCAGCATTTT 4014

RESULT 2
US-10-133-172-3
; Sequence 3, Application US/10133172
; Publication No. US20030104987A1
; GENERAL INFORMATION:
; APPLICANT: WEIGEL, PAUL H
; APPLICANT: WEIGEL, JANET A
; TITLE OF INVENTION: METHODS OF USING THE HYALURONAN RECEPTOR FOR ENDOCYTOSIS
; FILE REFERENCE: 5864.014
; CURRENT APPLICATION NUMBER: US/10/133,172
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 60/286,468
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 09/842,930
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 4642
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-133-172-3

Alignment Scores:
Pred. No.: 3,81e-231 Length: 4642
Score: 1772.00 Matches: 334
Percent Similarity: 95.18% Conservative: 2
Best Local Similarity: 94.62% Mismatches: 17
Query Match: 95.01% Indels: 0
DB: 14 Gaps: 0

US-09-466-778B-11 (1-353) x US-10-133-172-3 (1-4642)

Qy 1 MetThrGlyProGlyLysHisCysGluCysLysSerHisTyrValGlyAspGlyLeu 20
Db 3022 ATGACAGCCCGGGCAACCAAGTGTGAGTGAAGTCACTATGTCGGAGATGGGCTG 3081

Qy 21 AsnCysGluProGluLeuProLeuAspArgCysLeuGlnAspAsnGlyGlnCysHis 40
Db 3082 AACTGTGACCGGAGCAGCTGCCATTCACCGCTGCTTACAGGACAATGGCAGTGCAT 3141

Qy 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60
Db 3142 GCAGACCCCAATGTGTCGACCTCCACTTCAGGATACCACTGTGGGGTGTTCATCTA 3201

Qy 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80
Db 3202 CGCTCCCACTGGGCCAGTATAGCTGACCTTTGACAAAGCCAGAGAGGCTGTGCCAAC 3261

Qy 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis 100
Db 3262 GAAAGTCCGACCATGGCAACCTACCAACCACTCTCTCTATGCCAGAGGCAAGTATCCAC 3321

Qy 101 LeuCysSerAlaGlyTyrLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120
Db 3322 CTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTGGCTTACCCACAGCCTTTCGCTCC 3381

Qy 121 GlnAsnCysGlySerGlyValValGlyIleValAspTyrGlyProArgProAsnLysSer 140
Db 3382 CAGACTGTGCTCTGTGTGTGGATAGTGGACTATGACCTAGACCTACACCAACAGAGT 3441

Qy 141 GluMetTyrAspValPheCysTyrArgMetLysAspValAsnCysThr***LysValGly 160
Db 3442 GAAATGTGGATGTCTTCTGCTATCGGATGAAAGATGTAAGTGAAGTGAAGTGAAGTGAAG 3501

Qy 161 TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuLeuGlnValLeuMetSerPhePro 180

Db 3502 TATGTGGAGATGGCTTCTCATGAGTGGGAGCTGCTGAGGTCCTGCTCC 3561
 Qy 181 SerLeuThrAnPheLeuThrGluValLeuAlaTyrSerAnSerAlaAtrGlyArg 200
 Db 3562 TCACTCAAACTTCTTGACGGAAGTGTGCTTATTCACAGCTCAGCTCGAGGCGT 3621
 Qy 201 AlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsn 220
 Db 3622 GCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGCACCTCTTTGTGCCACAGAAC 3681
 Qy 221 SerGlyLeuGlyGluAsnGluThrIleuSerGlyArgAspIleGluHisLeuAlaAsn 240
 Db 3682 AGTGGGCTGGGGAGATGAGACCTTGTCTGGGGGAGACATCGAGCACCACTCGCCAAAT 3741
 Qy 241 ValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGly 260
 Db 3742 GTACAGCATGTTTCTACATGACCTTGTCAATGGCACCCCTGCAACGAGGCTGGGA 3801
 Qy 261 SerLysLeuLeuIleThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal 280
 Db 3802 AGCAAGCTGCTCATCACTGCGCAGCAGGACCACTCCCAACCGAGAGACCAAGTTGTT 3861
 Qy 281 AspGlyArgAspThrLeuGluThrAspIleCysAlaSerAsnGlyIleThrHisValle 300
 Db 3862 GATGGAGAGACCATTCGTGAGTGGGACATCTTTGCTCCATCCCAATGGGATCATTCATGTCT 3921
 Qy 301 SerArg***LeuLysAlaProProAlaProValThrLeu***HisThrGlyLeuGly*** 320
 Db 3922 TCCAGGCTTTAAAGACACCCCTGCCCCGTGACCTTGACCCACACTGGCTTGGGAGCA 3981
 Qy 321 GlyIlePhe*****IleIleLeuValThrGlyAlaValAlaLeuAlaTyrSerTyr 340
 Db 3982 GGGATCTTCTTGGCATCATCTGCTGACTGGGCTGTGCTTGGCTTGGCTTACTCTCTAC 4041
 Qy 341 PheArgIleAsnArgLysThrIleGlyPhe***HisPhe 353
 Db 4042 TTTCCGATAAACCGGAGAACAACTCGGCTTCCAGCATTTT 4080

RESULT 3

US-10-133-172-19
 ; Sequence 19, Application US/10133172
 ; Publication No. US20030104987A1
 ; GENERAL INFORMATION:
 ; APPLICANT: WEIGEL, PAUL H
 ; APPLICANT: WEIGEL, JANET A
 ; TITLE OF INVENTION: METHODS OF USING THE HYALURONAN RECEPTOR FOR ENDOCYTOSIS
 ; FILE REFERENCE: 5864.014
 ; CURRENT APPLICATION NUMBER: US/10/133.172
 ; PRIOR FILING DATE: 2002-04-25
 ; PRIOR APPLICATION NUMBER: 60/286,468
 ; PRIOR FILING DATE: 2001-04-25
 ; PRIOR APPLICATION NUMBER: 09/842,930
 ; PRIOR FILING DATE: 2001-04-25
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 19
 ; LENGTH: 4962
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-133-172-19

Alignment Scores:

Pred. No.:	4.24e-231	Length:	4962
Score:	1772.00	Matches:	334
Percent Similarity:	95.18%	Conservative:	2
Best Local Similarity:	94.62%	Mismatches:	17
Query Match:	95.01%	Indels:	0
DB:	14	Gaps:	0

US-09-466-778B-11 (1-353) x US-10-133-172-19 (1-4962)

Qy 1 MetThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeu 20

Db 3733 ATGACAGGCCCGCGCAAGCAAGTGTGAGTGTAAAGTCACTATGTCGAGATGGGCTG 3792
 Qy 21 AsnCysGluProGluGlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis 40
 Db 3793 AACTGTGAGCGGAGCGAGCTGCCCATGTGACCGTGTCTTACAGGACAAATGGGCGAGTGCAT 3852
 Qy 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60
 Db 3853 GCAGAGCCCAATGTGTGCGACCTCCACTCCAGGATACCACTGTGGGGTGTTCATCTA 3912
 Qy 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80
 Db 3913 CGTCCCCACTGGGCCAGTATAGCTGACCTTTGACAAAGCCAGAGAGGCTGTGCCAAC 3972
 Qy 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis 100
 Db 3973 GAGCTGGGACCAATGGCACTTACCACTGCTCTCTATGCCAGAGGCCAAGTACCAC 4032
 Qy 101 LeuCysSerAlaGlyTrpLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120
 Db 4033 CTGTGCTCAGAGGCTGGCTGGAGACCGGGCGGGTTCCTTACCCACAGCCTTCGCCCTC 4092
 Qy 121 GlnAsnCysGlySerGlyValValGlyIleValAspTyrGlyProArgProAsnLysSer 140
 Db 4093 CAGAACTGTGGCTCTGCTGTGGTGGGATAGTGGACTATGGACCTAGACCCCAACAGAGT 4152
 Qy 141 GluMetTrpAspValPheCysTyrArgMetLysAspValAsnCysThr***LysValGly 160
 Db 4153 GAAATGTGGGATGTCTTCTGCTATCGATGAAAGATGTGAATGACCTGCAGGTGGG 4212
 Qy 161 TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuGlnValLeuMetSerPhePro 180
 Db 4213 TATGTGGAGATGGCTTCTCATGAGTGGGAACTGCTGAGGTCTGTGTCTTCTTCCCC 4272
 Qy 181 SerLeuThrAnPheLeuThrGluValLeuAlaTyrSerAnSerSerAlaAtrGlyArg 200
 Db 4273 TCACCTCAAACTTCTCGAGGAGTGTGCTGCTTATTCCAACAGCTCAGCTCGAGGCGGT 4332
 Qy 201 AlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsn 220
 Db 4333 GCATTTCTAGAACACCTGACTGCTGCTCATCGCGGCGACCTCTTTGTGGCCACAGAAC 4392
 Qy 221 SerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisLeuAlaAsn 240
 Db 4393 AGTGGGCTGGGGAGAAATGAGACCTTGTCTGGGGGAGACATCGAGCACCACTCGCCAAT 4452
 Qy 241 ValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGly 260
 Db 4453 GTACAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCCCTTGCACCAACGAGGCTGGGA 4512
 Qy 261 SerLysLeuLeuIleThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal 280
 Db 4513 AGCAAGCTGTCTCATCACTGCGCAGCAGGACCCACTCCCAACGAGGAGACCAAGTTGTT 4572
 Qy 281 AspGlyArgAspThrLeuGluThrAspIleCysAlaSerAsnGlyIleThrHisValle 300
 Db 4573 GATGGAAGAGCAATTCGTGAGTGGGACATCTTTGCCCTCCAATGGGATCATTCATGTCAAT 4632
 Qy 301 SerArg***LeuLysAlaProProAlaProValThrLeu***HisThrGlyLeuGly*** 320
 Db 4633 TCCAGGCTTTAAAGACACCCCTTGGCCCCGTGACCTTGACCTTGACCCACACTGGCTGGGAGCA 4692
 Qy 321 GlyIlePhe*****IleIleLeuValThrGlyAlaValAlaLeuAlaTyrSerTyr 340
 Db 4693 GGGATCTTCTTGGCATCATCTGCTGGTGGTGGGCTGTGCTTGGCTTGGCTTACTCTCTAC 4752
 Qy 341 PheArgIleAsnArgLysThrIleGlyPhe***HisPhe 353
 Db 4753 TTTCCGATAAACCGGAGAACAACTCGGCTTCCAGCATTTT 4791

RESULT 4

US-10-028-248A-210

```

; Sequence 210, Application US/10028248A
; Publication No. US20030235882A1
; GENERAL INFORMATION:
; APPLICANT: Shimketa, Richard
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corine
; APPLICANT: Casman, Stacie
; APPLICANT: Malyankar, Uriel
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Gangolli, Esha
; APPLICANT: Miller, Charles
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Smithson, Glennnda
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Liu, Xiaohong
; APPLICANT: Colman, Steven
; APPLICANT: Tchernev, Velizar
; APPLICANT: Si, Jingsheng
; APPLICANT: Edinger, Shlomit
; APPLICANT: Stone, David
; APPLICANT: Scioe, Paul
; APPLICANT: Millet, Isabelle
; APPLICANT: Rothenberg, Mark
; TITLE OF INVENTION: No. US20030235882A1el Nucleic Acids and Polypeptides and Methods
; FILE OF INVENTION: Thereof
; FILE REFERENCE: 21402-222
; CURRENT APPLICATION NUMBER: US/10/028,248A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256619
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262959
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311266
; PRIOR FILING DATE: 2001-08-09
; NUMBER OF SEQ ID NOS: 211
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 210
; LENGTH: 3260
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3171)
; OTHER INFORMATION: Wherein n is A, C, G, or T
US-10-028-248A-210

Alignment Scores:
Pred. No.: 5,58e-231 Length: 3260
Score: 1769.00 Matches: 333
Percent Similarity: 95.18% Conservative: 3
Best Local Similarity: 94.33% Mismatches: 17
Query Match: 94.85% Indels: 0
DB: 15 Gaps: 0

US-09-466-778B-11 (1-353) x US-10-028-248A-210 (1-3260)
QY 1 MetThrGlyProGlyLysHisLysCysGluCysLysSerHisThrValGlyAspGlyLeu 20
DB 1635 ATGACAGCGCCGGGCAAGCACAGGTGTGAGTGTAAAGTCACTATGTGGAGATGGGCTG 1694
QY 21 AsnCysGluProGluGlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis 40
DB 1695 AACTGTGACCGGAGCAGCTGCCCATTTACCGCTGTACAGACACATGGGCGATGCCAT 1754

```

```

QY 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60
DB 1755 GCAGAGCCCAATGTGTGACCTCCACTTCAGGATACCACTGTGGGTGTTCATCTA 1814
QY 61 ArgSerProLeuGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80
DB 1815 CGTCCCTCCACTGGGCCAGTATAAGCTTGTGACAAAGCCAGAGAGGCTGTGCCAAC 1874
QY 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis 100
DB 1875 GAAGCTCGGACCATGGCACTCAACAGCTCTCTATGCCAGAGGCCAAGTACCAC 1934
QY 101 LeuCysSerAlaGlyTyrLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120
DB 1935 CTGTGCTCAGCAGGCTGGCTGGAGACGGGGTTCCTACCCACAGAGCTTCGGCTCC 1994
QY 121 GlnAsnCysGlySerGlyValValGlyLysValAspTyrGlyProArgProAsnLysSer 140
DB 1995 CAGAACTGTGGCTCTGGTGTGGTTGGATAGTGGACTATGGACCTAGACCCACAGAGT 2054
QY 141 GluMetTrpAspValPheCysTyrArgMetLysAspValAsnCysThr***LysValGly 160
DB 2055 GAAATGTGGATGTCTTCGTATCGATGAAAGATGTGAATGCACCTGCAAGGTGGC 2114
QY 161 TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuGlnValLeuMetSerPhePro 180
DB 2115 TATGTGGAGATGGCTTCTCATGCAGTGGGAACCTGTGCAGGTCTGTATGCTCTCCC 2174
QY 181 SerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArg 200
DB 2175 TCACCTCAAACTTCTTCAGCGAAGTGTGGCTATTTCCACAGCTCAGCTCGAGGCGGT 2234
QY 201 AlaPheLeuGlnHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsn 220
DB 2235 GCATTTCTAGAACACCTGACTGACCTGTCATCGCGGCACCTCTTTGTGCCACAGAAC 2294
QY 221 SerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisLeuAlaAsn 240
DB 2295 AGTGGCTGGGGGAGATGAGACCTGTCTGGCGGGACATCGAGACCACTCGCCAAT 2354
QY 241 ValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGly 260
DB 2355 GTGAGCATGTTTTCTACATGACCTTGTCAATGGCACCCCTTGCAACAGAGGGTGGGA 2414
QY 261 SerLysLeuLeuIleThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal 280
DB 2415 AGCAAGCTGTCTCATCTGCGCAGCCAGCACCTCCAAACCGACGAGACCAAGTGTGT 2474
QY 281 AspGlyArgAspThrLeuGluTyrAspIleCysAlaSerAsnGlyIleThrHisValIle 300
DB 2475 GATGGAAGAGCCATTTGCGATGGGACATCTTGCCTCCAAATGGGATCATTCATGTCAAT 2534
QY 301 SerArg***LeuLysAlaProProAlaProValThrLeu***HisThrGlyLeuGly*** 320
DB 2535 TCCAGGCGCTTTAAAGCACCCCTGCCCCGTGACCTTGACCCACACTGGCTTGGGAGCA 2594
QY 321 GlyIlePhe*****IleLeuValThrGlyAlaValAlaLeuAlaAlaTyrSerTyr 340
DB 2595 GGGATCTTCTTTGGCATCATCTGCTGACTGGGGGTGTGTGCTTGGTGTCTTACTCTTAC 2654
QY 341 PheArgIleAsnArgLysThrIleGlyPhe***HisPhe 353
DB 2655 TTTCGATAAAACCGGAGAACCAATCGGCTTCCAGCATTTT 2693

```

```

RESULT 5
US-10-107-782-210
; Sequence 210, Application US/10107782
; Publication No. US20040018970A1
; GENERAL INFORMATION:
; APPLICANT: Boldog, Ferenc,
; APPLICANT: Casman, Stacie,
; APPLICANT: Colman, Steve,

```

```

; APPLICANT: Edinger, Shlomit,
; APPLICANT: Gangolli, Baha,
; APPLICANT: Kekuda, Ramesh,
; APPLICANT: Li, Li,
; APPLICANT: Liu, Xiaohong,
; APPLICANT: Malyankar, Uriel,
; APPLICANT: Miller, Charles,
; APPLICANT: Millet, Isabelle,
; APPLICANT: Patturajan, Meera,
; APPLICANT: Rothenberg, Mark,
; APPLICANT: Sciore, Paul,
; APPLICANT: Shenoy, Suresh,
; APPLICANT: Shimkets, Richard,
; APPLICANT: Si, Jingsheng,
; APPLICANT: Smithson, Glennda,
; APPLICANT: Spytek, Kimberly,
; APPLICANT: Stone, David,
; APPLICANT: Taupier, Raymond, Jr.,
; APPLICANT: Tchernev, Velizar,
; APPLICANT: Vernet, Corine,
; APPLICANT: Zerhusen, Brian
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES AND METHODS OF USE THEREOF
; FILE INVENTION: 21402-222CIP
; CURRENT APPLICATION NUMBER: US/10/107,782
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: 10/028,248
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,619
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262,959
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272,408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285,189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308,039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311,266
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/279,344
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 215
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 210
; LENGTH: 3260
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (171)..(2861)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3171)..(3171)
; OTHER INFORMATION: Wherein n may be a, c, g or t
US-10-107-782-210

```

```

Alignment Scores:
Pred. No.: 5.58e-231 Length: 3260
Score: 1769.00 Matches: 333
Percent Similarity: 95.18% Conservative: 3
Best Local Similarity: 94.33% Mismatches: 17
Query Match: 94.85% Indels: 0
DB: 15 Gaps: 0

US-09-466-778b-11 (1-353) x US-10-107-782-210 (1-3260)

```

```

Qy 1 MetThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeu 20
Db 1635 ATGACAGCCCGGCGGACGACCAAGTGTGAGTGTAAAGTCACTATGTCGGAGATGGGCTG 1694
Qy 21 AsnCysGluProGluGlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis 40
Db 1695 AACTGTGAGCGGAGCAGCTGCCCATTTGACCGCTGTACAGGACATGGCAGTGGCAT 1754

```

```

Qy 41 AlaAspAlaLysCysValAspLeuHisPheGluAspThrThrValGlyValPheHisLeu 60
Db 1755 GCAGAGCCCAATGTGTGACCTCCACTTCCAGATACCACTGTGGGGTGTTCATCTA 1814
Qy 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80
Db 1815 CGCTCCCACTGGGCGAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCTGTGCCAAC 1874
Qy 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis 100
Db 1875 GAAGCTGGACCATGGCAACCTACACAGCTCTCTATGCCAGAGGCGCAAGTACCAC 1934
Qy 101 LeuCysSerAlaGlyTrpLeuGluThrGlyArgValAlaLysTyrProThrAlaPheAlaSer 120
Db 1935 CTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTGTCCCTACCCCAAGCAGCTTC 1994
Qy 121 GlnAsnCysGlySerGlyValValGlyIleValAspTyrGlyProArgProAsnLysSer 140
Db 1995 CAGAACTGTGGCTCTGTGTGGTGGATAGTAGCTATGAGCTAGACCCCAAGAGT 2054
Qy 141 GluMetTrpAspValPheCysTyrArgMetLysAspValAsnCysThr***LysValGly 160
Db 2055 GAAATGTGGGATGCTCTCTGCTATCGGATGAAAGATGTGAAGTGCACCTGCAAGTGGG 2114
Qy 161 TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuGlnValLeuMetSerPhePro 180
Db 2115 TATGTGGGAGATGGCTTCTCATGAGTGGGAACCTGTGTCAGGTCTCTGTATGCTTCCC 2174
Qy 181 SerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArg 200
Db 2175 TCACCTCACAACTTCTGACGGAAGTGTGGCTATTCACACAGCTCAGCTCAGGCGCT 2234
Qy 201 AlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsn 220
Db 2235 GCATTTCTAGACACACCTGACTGACCTGTCCATCGGGGCCACCTCTTTGTGCCACAGAA 2294
Qy 221 SerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisLysLeuAlaAsn 240
Db 2295 AGTGGGCTGGGGAGAAATGAGACCTTGTCTGGCGGGGACATCGAGCACCACCTCGCCA 2354
Qy 241 ValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGly 260
Db 2355 GTCAGCATGTTTTTCTACAAATGACCTTGTCAATGGCAGCACCCTCGCAACAGGGTGGGA 2414
Qy 261 SerLysLeuLeuIleThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal 280
Db 2415 AGCAAGCTGCTCATCTACTGCCAGCAGGACCCACTCCAAACCGAGGAGACAGGTTGTT 2474
Qy 281 AspGlyArgAspThrLeuGluTrpAspIleCysAlaSerAsnGlyIleThrHisValIle 300
Db 2475 GATGGAGAGAGCCATTCTGCAAGTGGGACATCTTTGCCCTCCAATGGGATCATTCATGCT 2534
Qy 301 SerArg***LeuLysAlaProProAlaProValThrLeu***HisThrGlyLeuGly*** 320
Db 2535 TCCAGGCGCTTAAAGACACCCCTGCCCCGTGACCTTGACCCACACACTGGCTGGGAGCA 2594
Qy 321 GlyIlePhe*****IleIleLeuValThrGlyAlaValAlaLeuAlaLysTyr 340
Db 2595 GGGATCTCTTTGCCATCATCTGCTGGTGTGAGTGGGGCTGTTCCTTGGCTGTCTTACTCTAC 2654
Qy 341 PheArgIleAsnArgLysThrIleGlyPhe***HisPhe 353
Db 2655 TTTCGGATAAACCGAGACCAATCGGCTTCCAGCATTTT 2693

```

```

RESULT 6
US-10-028-248A-3
; Sequence 3, Application US/10028248A
; Publication No. US20030235892A1
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corine

```

```

; APPLICANT: Casman, Stacie
; APPLICANT: Malyankar, Uriel
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Gangolli, Esha
; APPLICANT: Miller, Charles
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Taupier Jr., Raymond J
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Smithson, Glennnda
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Liu, Xiaohong
; APPLICANT: Colman, Steven
; APPLICANT: Tchernev, Velizar
; APPLICANT: Si, Jingsheng
; APPLICANT: Edinger, Shlomit
; APPLICANT: Stone, David
; APPLICANT: Sciore, Paul
; APPLICANT: Millet, Isabelle
; APPLICANT: Rothenberg, Mark
; TITLE OF INVENTION: No. US20030235882A1el Nucleic Acids and Polypeptides and Methods
; FILE REFERENCE: 21402-222
; CURRENT APPLICATION NUMBER: US/10/028,248A
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256619
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262959
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311266
; PRIOR FILING DATE: 2001-08-09
; NUMBER OF SEQ ID NOS: 211
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (131)..(1316)
; OTHER INFORMATION: Wherein n is A, C, G, or T
US-10-028-248A-3

Alignment Scores:
Pred. No.: 2,13e-201 Length: 8495
Score: 1558.00 Matches: 300
Percent Similarity: 85.31% Conservative: 2
Best Local Similarity: 84.75% Mismatches: 15
Query Match: 82.54% Indels: 37
DB: 15 Gaps: 1

US-09-466-778B-11 (1-353) x US-10-028-248A-3 (1-8495)

QY 1 MetThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeu 20
Db 6948 ATGACAGGCGCGGCAAGCAAGTGTGAGTGAAGTCACTATGTCGAGATGGGTG 7007

QY 21 AsnCysGluProGluGlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis 40
Db 7008 AACTGTGAGCGGAGCAGCTGCCATTCACCGCTGCTTACAGACAAATGGGCGAGTGCAT 7067

QY 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrValGlyValPheHisLeu 60
Db 7068 GCAGACGCCAAATGTGTGACCTCCACTTCCAGGATACCACTGTGGGTGTTCATCTA 7127

QY 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80

```

```

Db 7128 CGTCCCACTGGGCCAGTATAGCTGACCTTTGACAAAGCCAGAGGCGCTGTGCCAAC 7187
QY 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis 100
Db 7188 GAGCTGGACCATGGCAACCTACACAGCTCTCTATGCCAGAGGCCAAGTACCAC 7247
QY 101 LeuCysSerAlaGlyTyrLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120
Db 7248 CTGTGCTCAGCAGGCTGGTGGAGACCGGGGGTGGCTACCCACAGCCCTTCGCCTCC 7307
QY 121 GlnAsnCysGlySerGlyValValGlyLysValAspTyrGlyProArgProAsnLysSer 140
Db 7308 CAGAACTGTGGCTCTGTGTGGTGGATAGTGACTATGACCTAGACCCCAACAGAGT 7367
QY 141 GluMetTyrAspValPheCysTyrArgMet-LysAspValAsnCysThr***LysValGI 160
Db 7368 GAATGTGGGATGCTTCTGCTATCGGATGAAAG----- 7401
QY 160 yTyrValGIAspGlyPheSerTyrSerGlyAsnLeuGlnValLeuMetSerPhePr 180
Db 7401 ----- 7401
QY 180 oSerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyAr 200
Db 7402 -----GAAGTGTGGCTATTTCCAAACAGCTCAGCTCAGGCGG 7439
QY 200 gAlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAs 220
Db 7440 TGCATTTCTAGAACACCTGACTGACCTGCTCCATCCGGGGACCCCTCTTTGTGCCACAGAA 7499
QY 220 nSerGlyLeuGluAsnGluThrLeuSerGlyArgAspIleGluHisLeuAlaAs 240
Db 7500 CAGTGGGCTGGGGAGAAATGAGACCTTGTCTGGGGGGGACATCGAGCACCCACCTCGCCAA 7559
QY 240 nValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGI 260
Db 7560 TGTGAGCATGTTTTTCTACAATGACCTGTGTCAATGGACACCCCTGCAAAACGAGCTGGG 7619
QY 260 ySerLysLeuLeuIleThrAspArgGlnAspProLeuHisProThrGluThrArgCysVa 280
Db 7620 AAGCAAGCTGCTCATCTGCTCCAGCAGGACCCACTCCACCGAGGAGCCAGGTTTGT 7679
QY 280 lAspGlyArgAspThrLeuGluTyrAspIleCysAlaSerAsnGlyIleThrHisValII 300
Db 7680 TGATGGAAGAGCCATTCTGCACTGGGACATCTTTGCCCTCCAATGGGATCATTCATGTCAT 7739
QY 300 eSerArg***LeuLysAlaProProAlaProValThrLeu***HisThrGlyLeuGly** 320
Db 7740 TTCAGGCGCTTAAAGACACCCCTGCCCGCTGACCTTGACCCACACTGGCTTGGGAGC 7799
QY 320 *GlyIlePhe*****IleIleLeuValThrGlyAlaValAlaLeuAlaLysTyrSerTy 340
Db 7800 AGGATCTTCTTGGCATCATCTGTTGACTGGGCTGTTGCCCTGGCTGTCTACTCTACTCTA 7859
QY 340 rPheArgIleAsnArgLysThrIleGlyPhe***HisPhe 353
Db 7860 CTTTGGGATAAACCGGAGACAAATCGGCTTCCAGCATTTT 7899

RESULT 7
US-10-107-782-3
; Sequence 3, Application US/10107782
; Publication No. US20040018970A1
; GENERAL INFORMATION:
; APPLICANT: Boldog, Ferenc,
; APPLICANT: Casman, Stacie
; APPLICANT: Colman, Steve
; APPLICANT: Edinger, Shlomit,
; APPLICANT: Gangolli, Esha,
; APPLICANT: Kekuda, Ramesh,
; APPLICANT: Li, Li,
; APPLICANT: Liu, Xiaohong,
; APPLICANT: Malyankar, Uriel,

```


Db 7680 TGATGGAAGACCAATCTCGATGGGACATCTTTCCTCCAAATGGGATCATTCATGTGCT 7739
 Qy 300 eSerArg**LeuLysAlaProAlaProAlaProValThrLeu**HisThrGlyLeuGly** 320
 Db 7740 TTCAGGCGCTTTAAAGACACCCCTGCCCCGTGACCTTGACCCACATCGCTTGGGAGC 7799
 Qy 320 *GlyIlePhe*****IleLeuValThrGlyAlaValAlaLeuAlaTyrSerTy 340
 Db 7800 AGGGATCTCTTTGGCATCATCTCGTGACTGGGCGCTGTGCTTGGCTTACTCCTA 7859
 Qy 340 rPheArgIleAsnArgLysThrIleGlyPhe***HisPhe 353
 Db 7860 CTTTCGGATAAACCGGAGACAATCGGCTTCAGCATTTT 7899

RESULT 8

US-10-028-248A-1
 ; Sequence 1, Application US/10028248A
 ; Publication No. US20030235882A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Shimkets, Richard
 ; APPLICANT: Patturajan, Meera
 ; APPLICANT: Vernet, Corine
 ; APPLICANT: Casman, Stacie
 ; APPLICANT: Malyankar, Uriel
 ; APPLICANT: Shenoy, Suresh
 ; APPLICANT: Spytek, Kimberly
 ; APPLICANT: Gangolli, Esha
 ; APPLICANT: Miller, Charles
 ; APPLICANT: Boldog, Ferenc
 ; APPLICANT: Li, Li
 ; APPLICANT: Taupier Jr, Raymond J
 ; APPLICANT: Kekuda, Ramesh
 ; APPLICANT: Smithson, Glennda
 ; APPLICANT: Zerhusen, Bryan
 ; APPLICANT: Liu, Xiaohong
 ; APPLICANT: Colman, Steven
 ; APPLICANT: Tchernev, Velizar
 ; APPLICANT: Si, Jingsheng
 ; APPLICANT: Edinger, Shlomit
 ; APPLICANT: Stone, David
 ; APPLICANT: Sciore, Paul
 ; APPLICANT: Millet, Isabelle
 ; APPLICANT: Rothenberg, Mark
 ; TITLE OF INVENTION: No. US20030235882A1el Nucleic Acids and Polypeptides and Methods
 ; FILE OF INVENTION: Thereof
 ; CURRENT APPLICATION NUMBER: US/10/028,248A
 ; CURRENT FILING DATE: 2001-12-19
 ; PRIOR APPLICATION NUMBER: 60/256619
 ; PRIOR FILING DATE: 2000-12-19
 ; PRIOR APPLICATION NUMBER: 60/262959
 ; PRIOR FILING DATE: 2001-01-19
 ; PRIOR APPLICATION NUMBER: 60/272408
 ; PRIOR FILING DATE: 2001-02-28
 ; PRIOR APPLICATION NUMBER: 60/285189
 ; PRIOR FILING DATE: 2001-04-20
 ; PRIOR APPLICATION NUMBER: 60/308039
 ; PRIOR FILING DATE: 2001-07-26
 ; PRIOR APPLICATION NUMBER: 60/311266
 ; PRIOR FILING DATE: 2001-08-09
 ; NUMBER OF SEQ ID NOS: 211
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 8444
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (163)..(168)
 ; OTHER INFORMATION: wherein n is A, C, G, or T
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1279)..(1284)

; OTHER INFORMATION: wherein n is A, C, G, or T
 US-10-028-248A-1
 Alignment Scores:
 Pred. NO.: 7.44e-201 Length: 8444
 Score: 1554.00 Matches: 296
 Percent Similarity: 85.27% Conservative: 5
 Best Local Similarity: 83.85% Mismatches: 16
 Query Match: 83.32% Indels: 36
 DB: 15 Gaps: 1
 US-09-466-778B-11 (1-353) x US-10-028-248A-1 (1-8444)
 Qy 1 MetThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeu 20
 Db 6907 ATGACAGGCCCGGGCAAGCAAGTGTGAGTGTAAAGTCACTATGTCGGAGATGGGCTG 6966
 Qy 21 AsnCysGluProGluGlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis 40
 Db 6967 AACTGTGAGCGGAGCAGCTGCCAATTGACCGCTGTACAGGACAATGGGACGTGCCAT 7026
 Qy 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60
 Db 7027 GCAGACGCCAATGTGTGCACCTCCACTTCCAGGATACCACTGTGGGGTGTCCACTA 7086
 Qy 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80
 Db 7087 CGCTCCCGACCTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCGCTGTCCAAC 7146
 Qy 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis 100
 Db 7147 GAAGTCGCACCATGGCAACCTACACAGCTCTCTCTATGCCCAAGAGGCCAAGTACCAAC 7206
 Qy 101 LeuCysSerAlaGlyTyrLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120
 Db 7207 CTGTGCTCAGCAGCGCTGGCTGGAGCCGGGGGTTGCCTACCCACAGCCCTTCGCTCC 7266
 Qy 121 GlnAsnCysGlySerGlyValValGlyIleValAspTyrGlyProArgProAsnLysSer 140
 Db 7267 CAGAACTGTGGCTCTGCTGTGGTGGATAGTGACCTATGACCTAGACCTAGACCCCAACAGAGT 7326
 Qy 141 GluMetTrpAspValPheCysTyrArgMetLysAspValAsnCysThr***LysValGly 160
 Db 7327 GAAATGTGGGATGTCTTCTGTCTATCGGATGAAG----- 7359
 Qy 161 TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuGlnValLeuMetSerPhePro 180
 Db 7359 ----- 7359
 Qy 181 SerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArg 200
 Db 7360 -----GAAAGTGTGGCTCTATTCCAAACAGCTCAGCTCGAGGCGGT 7398
 Qy 201 AlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsn 220
 Db 7399 GCATTTCTAGAACACCTGACTGACCTGTCCATCCGGGACCCCTTTGTGTCACAGAAC 7458
 Qy 221 SerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisLeuAlaAsn 240
 Db 7459 AGTGGCTGGGGAGAAATGAGACCTTGTCTGGGGGGGACATCGAGCACCCACCTCGCCAAT 7518
 Qy 241 ValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGly 260
 Db 7519 GTCAGCATGTTTTTCTCAATGACCTTGTCAATGGCACCCCTCGCAACAGAGGCTGGGA 7578
 Qy 261 SerLysLeuLeuIleThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal 280
 Db 7579 AGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGGTACAAAGTAGTGTGTT 7638
 Qy 281 AspGlyArgAspThrLeuGluThrAspIleCysAlaSerAsnGlyIleThrHisValIle 300
 Db 7639 GATGGAAGAGCCATCTTCAGTGGGACATCTTTCCTCCCAATGGGATCATTCATGTCATT 7698

Qy	301	SerArg*	LeuLysAlaProProAlaProValThrLeu***HisThrGlyLeuGly***	320
Db	7699	TCAGGCGCTTTAA	AAAGACACCCCTGCCCCGTGACCTGACCCACACTGGCTTGGAGCA	7758
Qy	321	GlyIlePhe*****Ile	LeuValThrGlyAlaValAlaLeuAlaIaIaTySerTyr	340
Db	7759	GGGATCTCTTTT	TGCATCATCTCGTGACTGGGGCTGTGCGTTGGTGCTTACTCTAC	7818
Qy	341	PheArgIleAsnArgLysThrIleGlyPhe***HisPhe	353	
Db	7819	TTTCGGATAAACCG	GAGAACAAATCGGTACCAGCATTTT	7857

RESULT 9

```

US-10-107-782-1
; Sequence 1, Application US/10107782
; Publication No. US20040018970A1
; GENERAL INFORMATION:
; APPLICANT: Boldog, Ferenc,
; APPLICANT: Casman, Stacie
; APPLICANT: Colman, Steve,
; APPLICANT: Edinger, Shlomit,
; APPLICANT: Gangolli, Esna,
; APPLICANT: Kekuda, Ramesh,
; APPLICANT: Li, Li,
; APPLICANT: Liu, Xiaohong,
; APPLICANT: Malyankar, Uriel,
; APPLICANT: Miller, Charles,
; APPLICANT: Millet, Isabelle,
; APPLICANT: Patturajan, Meera,
; APPLICANT: Rothenberg, Mark,
; APPLICANT: Sciore, Paul,
; APPLICANT: Shenoy, Suresh,
; APPLICANT: Shinkets, Richard,
; APPLICANT: Si, Jingsheng,
; APPLICANT: Smithson, Glennda,
; APPLICANT: Spytek, Kimberly,
; APPLICANT: Stone, David,
; APPLICANT: Taupier, Raymond, jr.,
; APPLICANT: Tchernev, Velizar,
; APPLICANT: Vernet, Corinne,
; APPLICANT: Zernhusen, Brian
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: 21402-222CIP
; CURRENT APPLICATION NUMBER: US/10/107,782
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: 10/028,248
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,619
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262,959
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272,408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285,189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308,039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311,266
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/279,344
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 215
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 1
; LENGTH: 8444
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(8025)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (163)..(163)

```

QY 1 MetThrGlyProGlyValHisCysGluCysLysSerHisTyrValGlyAspGlyLeu 20
DB 6907 ATGACAGGCCCGGCAAGCACAAGTGTGAGTGTAAAGTCACTATGTGGAGATGGGTG 6966
QY 21 AsnCysGluProGluGlnLeuProLeuAspArgCysLeuGlnAspAsnGlyLysHis 40
DB 6967 AACTGTGACCGGAGAGAGTGCCTTACCCCTGCTTACAGGACAATGGGAGTGCCTAT 7025
QY 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrValGlyValPheHisLeu 60
DB 7027 GCAGAGCGCAATGTGTGACCTCCACTTCCAGGATACCACTGTGGGTGTTCATCTA 7086
QY 61 ArgSerProLeuGlyClnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80
DB 7087 CGCTCCCACTGGGCGAGTAAAGTACCTTTGCAAAAGCCAGAGAGCCCTGTGCCAAC 7146
QY 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis 100
DB 7147 GAAGCTGGACCATGGCAACCTACAAACAGCTCTCTATGCCAGAGCCCAAGTACCAC 7206
QY 101 LeuCysSerAlaGlyTyrLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120
DB 7207 CTGTGCTCAGCAGGCTGGTGGAGACCGGGCGGTGCTTACCCACAGCCCTTCGCTCC 7266
QY 121 GlnAsnCysGlySerGlyValValGlyLysValAspTyrGlyProArgProAsnLysSer 140
DB 7267 CAGAACTGTGGCTGTGTGTGGGTAGTGGACTATGGACCTAGAGCCCAACAGAGT 7326
QY 141 GluMetTyrAspValPheCysTyrArgMetLysAspValAsnCysThr***LysValGly 160
DB 7327 GAAATGTGGATGTCTTCTGCTATCGGATGAAG----- 7359
QY 161 TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuLeuGlnValLeuMetSerPhePro 180
DB 7359 ----- 7359
QY 181 SerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArg 200
DB 7360 -----GAAGTGTGGCTTATTCACAGCTCAGCTCGAGCCGT 7398
QY 201 AlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsn 220
DB 7399 GCATTTCTAGAACACCTGACTGACCTGCTCATCCGCGGACCCCTCTTTGTGCCACAGAAC 7458
QY 221 SerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisLeuAlaAsn 240
DB 7459 AGTGGCTGGGGAGAAATGAGACCTTGTCTGGCGGGACATCGAGCACCACCTCGCCAA 7518
QY 241 ValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGly 260
DB 7519 GTCAGCATGTTTTCTACAAATGACCTTGTCAATGGCACCCCTGCAACAGAGGTGGGA 7578
QY 261 SerLysLeuLeuIleThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal 280
DB 7579 AGCAAGCTCTCATCTACCTCCAGCAGGACCCACTCCAAACCGGTACAAAGTAGTGTGT 7638
QY 281 AspGlyArgAspThrLeuGluThrAspIleCysAlaSerAsnGlyIleThrHisValIle 300
DB 7639 GATGGAAGAGCCATTCTCGAGTGGGACATCTTGGCTCCAAATGGGATCATTCATGTCA 7698
QY 301 SerArg***LeuLysAlaProAlaProValThrLeu***HisThrGlyLeuGly*** 320
DB 7699 TCCAGGCTTTAAAGACACCCCTGCCCGCTGACCTTGACCCACACTGGCTGGGAGCA 7758
QY 321 GlyIlePhe*****IleIleLeuValThrGlyAlaValAlaLeuAlaIleThrSerTyr 340
DB 7759 GGGATCTTCTTTGTCATCTCTGTGTGACTGGGCTGTGGCTTGGCTTGGCTTCTTACTCCTAC 7818
QY 341 PheArgIleAsnArgLysThrIleGlyPhe***HisPhe 353
DB 7819 TTTCGGATAAACCGGAGAACATCGGCTACCGACATTTT 7857

RESULT 10

US-09-842-930A-1
; Sequence 1, Application US/09842930A
; Publication No. US20020197681A1
; GENERAL INFORMATION:
; APPLICANT: Weigel, Paul
; TITLE OF INVENTION: Identification of Hyaluronan Receptor for Endocytosis
; FILE REFERENCE: 5820.503
; CURRENT APPLICATION NUMBER: US/09/842.930A
; CURRENT FILING DATE: 2001-04-22
; PRIOR APPLICATION NUMBER: 60/245,320
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: 60/199,538
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 4706
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-09-842-930A-1
Alignment Scores:
Pred. No.: 7,57e-196 Length: 4706
Score: 1514.50 Matches: 280
Percent Similarity: 87.54% Conservative: 29
Best Local Similarity: 79.32% Mismatches: 43
Query Match: 81.21% Indels: 1
DB: 9 Gaps: 1
US-09-466-778B-11 (1-353) x US-09-842-930A-1 (1-4706)
QY 1 MetThrGlyProGlyValHisCysGluCysLysSerHisTyrValGlyAspGlyLeu 20
DB 3022 ATGACGGGCCCGGCAAGCACAAGTGTGAAATGTAAGTCACTATGTGGGACGAGTG 3081
QY 21 AsnCysGluProGluGlnLeuProLeuAspArgCysLeuGlnAspAsnGlyLysHis 40
DB 3082 GACTGTGAGCTTGAGCAGCTGCCGCTCCAGCTTGTCTACAGGACAACGGACAGTGCAC 3141
QY 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrValGlyValPheHisLeu 60
DB 3142 CCAGATGCCAGCTGTGCAGACCTCTACTTCCAGGACACGACCTAGGAGTATTCATCTA 3201
QY 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80
DB 3202 CGCTCCCACTGGGCGAGTCAAACTGACATTTGCAAAAGCCAAAGAGCCTGTGCCAAA 3261
QY 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis 100
DB 3262 GAAGCTGGACCATAGCCACCTACACAGCTCTCTATGCCAGAGCCCAAGTATCAC 3321
QY 101 LeuCysSerAlaGlyTyrLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120
DB 3322 CTGTGCTCGGCGGTGGCTGGAGAGTGGCGGGTGTGCTTACCGACTACGCTATGCTCT 3381
QY 121 GlnAsnCysGlySerGlyValValGlyLysValAspTyrGlyProArgProAsnLysSer 140
DB 3382 CAGAACTGTGTGCAAACTTGTGGGATCTGACTACGATCCAGGATCCAGGCGCCACAGAGT 3441
QY 141 GluMetTyrAspValPheCysTyrArgMetLysAspValAsnCysThr***LysValGly 160
DB 3442 GAAATGTGGATGTCTTCTGTTACCGGATGAAGATGTGAATGCACCTGCAGCAGCAGC 3501
QY 161 TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuLeuGlnValLeuMetSerPhePro 180
DB 3502 TATGTGGGAGATGGCTTCTCGTCAGTGGGAACCTGCTGTCAGGCTCTCATGCTTCC 3561
QY 181 SerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArg 200
DB 3562 TCGCTCACAAATCTCTCCTGACAGAGTGTGGCTTTTCCAGAGCTCAGCCCGGAGCAG 3621
QY 201 AlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsn 220

3622	Db	GCATTTTGAACACCTGACTGACTGTCATCGGTGGCACCCCTGTTGTGGCCACAGAAC	3681
221	QY	SerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisHisLeuAlaAsn	240
3682	Db	AGTGGGCTACCGGGAAATAAGAGCCTCTCTGGCGGGACATTGAGCACCACCTCACTAAT	3741
241	QY	ValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGly	260
3742	Db	GTCAACGCTCCCTTTTACAATGACCTTGTCAATGGTATCCCTTCGTAGGAGTATGCTGGGA	3801
261	QY	SerLeuLeuLeuIleThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal	280
3802	Db	AGCCAACCTGCTCATTAACCTTCAGCCAGGACCACTCCAC---CAAGAGACCAAGTTTG	3958
281	QY	AspGlyArgAspThrLeuGluThrAspIleCysAlaSerAsnGlyIleThrHisValle	300
3859	Db	GATGGAAGATCCATCTTCGAGTCGGGACATCATGCGCGCAATGAAATCTCCATATTAT	3918
301	QY	SerArg***LeuLysAlaProProAlaProValThrLeu***HisThrGlyLeuGly***	320
3919	Db	TCGTGAACCTTTGAGAGCTCTCCACCGCAGCACCGCTGCCACTCTGCCCTGGGGACA	3978
321	QY	GlyIlePhe*****IleIleLeuValThrGlyAlaValAlaLeuAlaIatYrSerTyr	340
3979	Db	GGTATATTCTGCGCGTCGCTGGTCACTGGTGGGATTGCTCTGGCAGCTTACTCTTAC	4038
341	QY	PheArgIleAsnArgLysThrIleGlyPhe***HisPhe	353
4039	Db	TTCCGGTCAAGCAGCGAAACCACTGGTTTCCAGCGTTT	4077

RESULT 11

```

US-103-172-1
; Sequence 1, Application US/10133172
; Publication No. US20030104987A1
; GENERAL INFORMATION:
; APPLICANT: WEIGEL, PAUL H
; APPLICANT: WEIGEL, JANET A
; TITLE OF INVENTION: METHODS OF USING THE HYALURONAN RECEPTOR FOR ENDOCYTOSIS
; FILE REFERENCE: 5864.014
; CURRENT APPLICATION NUMBER: US/10/133,172
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 60/286,458
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 09/842,930
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 4706
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-103-172-1

```

Alignment Scores:	
Pred. No.:	7.57e-196
Score:	1514.50
Percent Similarity:	87.54%
Best Local Similarity:	79.32%
Query Match:	81.21%
DB:	14
	Gaps:
	Indels:
	Mismatches:
	Conservative:
	Matches:
	Length:

US-09-466-778B-11 (1-353) x US-10-133-172-1 (1-4706)

Qy	1	MerThrGlyProGlyLysHisLysCysGluCysLysSerHisIYrValGlyAspGlyLeu	20
	
Db	3022	ATGACGGCCAGCAGCATAAGTGTGAATGTAAGAGTCACATATGTCGGGACGAGGTG	3081
Qy	21	AsnCysGluProGluGlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis	40
	
Db	3082	GACTGTAGCGCTGAGCAGCTGCCGCTCGACGTTGCTTACAGACACAGACAGTCGCGAC	3141
Qy	41	AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu	60
	

3142	CCAGATGCCAGCTGTGCGAGACCTCTACTTCCAGGACACGACCGCTAGAGATTCCATCTTA	3201
QY	61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn	80
Db	3202 CGCTCCCCACTCGGCGCAGTACAAACTGACATTGTGACAAAGCCAAAGAGCCTGTGCCAAA	3261
QY	81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis	100
Db	3262 GAAGCTCGCACCATTACGCCACTCAACACGAGCTCTCTCTATGCCAAGAGGCCAAGATTCAC	3321
QY	101 LeuCysSerAlaGlyTrpLeuGluThrGlyAArgValAlaTyrProThrAlaPheAlaSer	120
Db	3322 CTGTGCTCGCGCGCTGCTGGAGAGTGGCGGGTTCCTTACCCGACTACGTATGGCTCT	3381
QY	121 GlnAsnCysGlySerGlyValValGlyIleValAspTyrGlyProArgProAsnLysSer	140
Db	3382 CAGAAGTGTGTGCAAAAGCTGTGTGGATCGTAGACTACGGATCCAGGGCCACACAGAGT	3441
QY	141 GluMetTrpAspValPheCysTyrArgMetLysAspValAsnCysThr***LysValGly	160
Db	3442 GAAATGTGGAGTCTCTCTGTATACCGGATGAAAGATGTGAACCTGCACCTCGAAGGCGAGGC	3501
QY	161 TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuGlnValLeuMetSerPhePro	180
Db	3502 TATGTGGAGATGGCTTCGTGCGATGGGAACCTGCTGCAGGTCTCATGTCTCTTCCCC	3561
QY	181 SerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArg	200
Db	3562 TCGCTCACAACTTCTCGACAGAGGTGTGGCTTTTTCAGAGCTTCAGCCCGAGACAG	3621
QY	201 AlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsn	220
Db	3622 GCATTTTGAACACACTGACTGACCTGTCCATCCGTGGCACCCCTGTTGTGCGCACAGAAC	3681
QY	221 SerGlyLeuGlyGlnAsnGluThrLeuSerGlyArgAspIleGluHisHisLeuAlaAsn	240
Db	3682 AGTGGGCTACCGGGAATAAGACCTGTCTGGCCGGGACATTGAGCACCACTCACTAAT	3741
QY	241 ValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGly	260
Db	3742 GTCAACGTCTCCTTTTACAAATGACCTTGTCAAATGGTACCTTTCTGAGGACTATGCTGGGA	3801
QY	261 SerLysLeuLeuIleThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal	280
Db	3802 AGCCAACCTGCTCATACCTTCAGCCAGGACCACTCCAC---CAAGAGACCAAGTTTGTG	3858
QY	281 AspGlyArgAspThrLeuGluTrpAspIleCysAlaSerAsnGlyIleThrHisValIle	300
Db	3859 GATGGGAATGCCATTCTGCAGTGGGACATCATCGCCGCCAATGGAATCCCTCCATATTATT	3918
QY	301 SerArg***LeuLysAlaProProAlaProValThrLeu***HisThrGlyLeuGly***	320
Db	3919 TCTGAACCTTTGAGAGCTCTCCCAAGGACCAACCGCTGCCACACTCTGCCCTGGGGACA	3978
QY	321 GlyIlePhe*****IleIleLeuValThrGlyValAlaValAlaLeuAlaIatYrSerTyr	340
Db	3979 GGTATATTCTGCGCGTGTCTGGTACGTGGTGGCATGTCTCTGGCAGCTTACTCTTAC	4038
QY	341 PheArgIleAsnArgLysThrIleGlyPhe***HisPhe	353
Db	4039 TTCGGCTTAAGCAGCAACCACTGGTTTTCACGCGTTT	4077

RESULT 12

```

RESULT 12
US-10-264-237-1320
; Sequence 1320, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCES: FA131PI
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450

```

```
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 1320
; LENGTH: 3681
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5)-(5)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (7)-(7)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9)-(9)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (31)-(31)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-264-237-1320

Alignment Scores:
Pred. No.: 4,046-79 Length: 3681
Score: 661.50 Matches: 145
Percent Similarity: 53.44% Conservative: 49
Best Local Similarity: 39.94% Mismatches: 148
Query Match: 35.47% Indels: 21
DB: 15 Gaps: 7

US-09-466-778B-11 (1-353) x US-10-264-237-1320 (1-3681)

Qy 2 ThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeuAsn 21
Db 2235 ACCGGCCTGAACACACACCGCGCTGTGAGTGCCACGAGGCTAGCTAGGCGATGGAGTGCAG 2294

Qy 22 Cys---GluProGluGlnLeuProLeuAspArgCysGluGlnAsnGlyGlnCysHis 40
Db 2295 TGTCTGGAGGAGTGGACACCGCTGTGACCGCTCTTGGGCGCAGCCCGCTGCGCAC 2354

Qy 41 AlaAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60
Db 2355 TCAGATGCCATGTGCACTGACCTGACCTTCCAGGAGAAACGGGCTGGCGTTTCCACCTC 2414

Qy 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80
Db 2415 CAGGCCACACGCGGCCCTTATGCTGAACTTTCGAGGCTGAGCGGCATGGGAACA 2474

Qy 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis 100
Db 2475 CAGGAGCGGCTGTGCTTCATCCCTCAGCTCTCTGCTGCCACGAGCTGGGCTTCCAC 2534

Qy 101 LeuCysSerAlaGlyTrpLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120
Db 2535 CTGTGCTCATGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2594

Qy 121 GlnAsnCysGlySerGlyValValGlyValGlyValGlyValGlyValGlyValGlyVal 140
Db 2595 GCGGATGTGGCAATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2654

Qy 141 GluMetTrpAspValPheCysTyrArgMetLysAspValAsnCysThr***LysValGly 160
Db 2655 GAACGCTGGATGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2714

Qy 161 TyrValGlyAspGlyPheSer---TyrSerGlyAsnLeuLeuGlnValLeuMetSerPhe 179
Db 2715 TTCGTGGGTGACGGGATGATGACGACGTCGATGGAAGCTGCTGATGTGCTGCTGCTGCTG 2774
```


Alignment Scores:
Pred. No.: 3,96e-77 Length: 2753
Score: 645.50 Matches: 143
Percent Similarity: 52.89% Conservative: 49
Best Local Similarity: 39.39% Mismatches: 150
Query Match: 34.61% Indels: 21
DB: 10 Gaps: 7

US-09-466-778B-11 (1-353) x US-09-969-730-13 (1-2753)

QY 2 ThrGlyProGlyLysHisCysGluCysLysSerHisTyrValGlyAspGlyLeuAsn 21
DB 264 ACCGCCCTGAACACACACCGCGCTGTGAGTCCACGAGGCTACGTAGGCGATGGAGCTGCAG 323

QY 22 Cys---GluProGluGlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis 40
DB 324 TGTCTGGAGGAGTCGAACACCTGTGACCGCTCTGGCCAGCCACCGCCCTGCCAC 383

QY 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60
DB 384 TCAGATGCATGTGMACTGCATCTCCAGGAGAAACGGCGCTGGCGTTTTCACCTC 443

QY 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80
DB 444 CAGGCCACAGCGGCCCTTGTGCTGAACCTTTCCGAGGCTGAGCGGCATCGCAAGCA 503

QY 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis 100
DB 504 CAGGAGCGCTCTGCTTCATCTCAGCTCTCTGCTGCCACGAGCTGGCGTTCCAC 563

QY 101 LeuCysSerAlaGlyTyrLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120
DB 564 CTGTGCTCATGGGTGGTGGCCAAATGGCTCCACTGCCACCTGTGTTTTCCTCTGT 623

QY 121 GlnAsnCysGlySerGlyValValGlyLysValAspTyrGlyProArgProAsnLysSer 140
DB 624 GCGGACTGTGGCAATGGTGGGTGGCTAGTACCTGGGTGGCGCCGCAAGAACTCTCA 683

QY 141 GluMetTyrAspValPheCysTyrArgMetLysAspValAsnCysThr***LysValGly 160
DB 684 GAAAGCTGGATGCTACTGCTCTGCTGGTGGCAAGATGGCTGGCGTGGCGAATGGC 743

QY 161 TyrValGlyAspGlyPheSer---TyrSerGlyAsnLeuGlnValLeuMetSerPhe 179
DB 744 TTCGTGGGTGACGGATCAGCACGTGCAATGGGAAGCTGCTGGATGTGCTGGCTGCCA 803

QY 180 ProSerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGly 199
DB 804 GCCAACTCTCCACCTTCTATGGGATGCTATTGGGCTATGCCAATGCCACCGCGGGT 863

QY 200 ArgAlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGln 219
DB 864 CTCGACTTCTGGATCTCTGGATGATGAGCTCAGCTATAGACACTCTTCTGCTCTGTC 923

QY 220 AsnSerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisLeuAla 239
DB 924 AATGAAGCTTTGTGGACAACTATGCTGTGAGTGGCCAGACTTGGAGTGCATGCCCTCC 983

QY 240 AsnValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeu 259
DB 984 AACGCCACCTCTTAAGTCCAC---GCCACCGAGGGAAGTTGTTCCGCCCCCACTCA 1040

QY 260 GlySerLysLeuLeuLeuThrAspArg-----GlnAspProLeuHisProThrGlu 276
DB 1041 GGCCTCAGCCTCATCATAGTCAGCGAGCGCCCTGACAAAGTTCCTGGGCCCTGTGGCC 1100

QY 277 ThrArgCysValAspGlyArgAspThrLeuGluTyrAspIleCysAlaSerAsnGlyIle 296
DB 1101 CCAGGACAGTTGTGGTTAGCCGTATCATTTGTGGGACATCATGCGCTTCAATGGCATC 1160

QY 297 ThrHisValIleSerArg***LeuLysAlaProAlaProValThrLeu----- 313

DB 1161 ATCATGCTCTGGCCAGCCCTCTCTGGCACCCCAACAGCCAGGCTGCTGGCGCT 1220

QY 314 -----***HisThrGlyLeuGly***GlyIlePhe*****IleIleLeuVal 329
DB 1221 GAAGCCCACTGTGGCGGAGGCTGGGGCT-----GTGCTTGCC 1262

QY 330 ThrGlyAlaVal-----AlaLeuAlaTyrSerTyrPheArgIleAsnArgLysThr 347
DB 1263 GCTGAGCACTGTGCTGGTGGTGGCGGAGCTCTACCTCGTGGCCGAGCAAGCC 1322

QY 348 IleGlyPhe 350
DB 1323 ATGGCTTT 1331

RESULT 15
US-10-621-363-13
; Sequence 13, Application US/10621363
; Publication No. US20040023283A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 90 Human Secreted Proteins
; FILE REFERENCE: P2013P2C1
; CURRENT APPLICATION NUMBER: US/10/621,363
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: 09/969,730
; PRIOR FILING DATE: 2001-10-06
; PRIOR APPLICATION NUMBER: 09/774,639
; PRIOR FILING DATE: 2001-02-01
; PRIOR APPLICATION NUMBER: 60/238,291
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 09/244,112
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: PCT/US98/16235
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: 60/056,371
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,732
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,366
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,364
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,370
; PRIOR FILING DATE: 1997-08-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 373
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 2753
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-621-363-13

Alignment Scores:
Pred. No.: 3,96e-77 Length: 2753
Score: 645.50 Matches: 143
Percent Similarity: 52.89% Conservative: 49
Best Local Similarity: 39.39% Mismatches: 150
Query Match: 34.61% Indels: 21
DB: 16 Gaps: 7

US-09-466-778B-11 (1-353) x US-10-621-363-13 (1-2753)

QY 2 ThrGlyProGlyLysHisCysGluCysLysSerHisTyrValGlyAspGlyLeuAsn 21
DB 264 ACCGCCCTGAACACACACCGCGCTGTGAGTCCACGAGGCTACGTAGGCGATGGAGCTGCAG 323

QY 22 Cys---GluProGluGlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis 40
DB 324 TGTCTGGAGGAGTCGAACACCTGTGACCGCTCTGGCCAGCCACCGCCCTGCCAC 383

QY 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60
DB 384 TGTCTGGAGGAGTCGAACACCTGTGACCGCTCTGGCCAGCCACCGCCCTGCCAC 383

Db 384 TCAGATGCCATGTGMACTGACCTGACCTCCAGGAGAAACGGGCTGGCGTTTCCACCTC 443
QY ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80
Db 444 CAGGCCACACAGCGCCCTATATGCTCTGAACCTTTTCGGAGGCTGAGCGGCTATCGAAGCA 503
QY GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis 100
Db 504 CAGGAGCGCTCCTCTGCTTATTCCTCAGCTCTCTGCTGCCAGCAGCTGGGCTTCCAC 563
QY 101 LeuCysSerAlaGlyTyrLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120
Db 564 CTGTGCTCATGGGCTGGCTGGCCATGGCTCCACTGCCACCTGTGGTTTCCCTGTG 623
QY 121 GlnAsnCysGlySerGlyValValGlyIleValAspTyrGlyProArgProAsnLysSer 140
Db 624 GCGGACTGTGGCAATGGTGGGTGGGCTAGTCAGCGCTGGGTGCCCGCAAGAACCTCTCA 683
QY 141 GluMetTyrPheValPheCysTyrArgMetLysValAspValAsnCysThr***LysValGly 160
Db 684 GAACGCTGGGATGCCTACTCTCCGTGTGCAAGATGTGGCTGCCGATGCCGAAATGGC 743
QY 161 TyrValGlyAspGlyPheSer---TyrSerGlyAsnLeuLeuGlnValLeuMetSerPhe 179
Db 744 TTCGTGGGTACGGGATCAGCACGTGCATGGGAAGCTGTGGATGTGCTGGGTGCCACT 803
QY 180 ProSerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGly 199
Db 804 GCCAACTTCTCCACCTTCTATGGGATGCTATTGGGCTATGCCAATGCCACCGCGGGT 863
QY 200 ArgAlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGln 219
Db 864 CTCGACTTCTGGACTTCTCGGATGTAGCTCACGTATAGACACTCTTCGTCCCTGTC 923
QY 220 AsnSerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisLeuAla 239
Db 924 AATGAGGCTTTGTGGACACATCACGCTGATGGGCCACAGCTTGGAGCTGCATGCCCTC 983
QY 240 AsnValSerMetPheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeu 259
Db 984 AACGCCACCTCTTAAGTGCAC---GCCAGCCAGGGGAAGTTGCTTCGGGCCCACTCA 1040
QY 260 GlySerLysLeuLeuIleThrAspArg-----GlnAspProLeuHisProThrGlu 276
Db 1041 GGCTCAGCCTCATCATCAGTACGACGACGAGGCCCTGACACAGTTCTCCGGGCCCTGTGGCC 1100
QY 277 ThrArgCysValAspGlyArgAspThrLeuGluTyrPheAspIleCysAlaSerAsnGlyIle 296
Db 1101 CCAGGGACAGTTGGTTAGCGGTATCATTTGTGGGACATCATGGCTTCATGGGATC 1160
QY 297 ThrHisValIleSerArg***LeuLysAlaProAlaProValThrLeu----- 313
Db 1161 ATCCATGCTCTGGCCAGCCCTCTCTGGCACCCCTCCAGCCACAGCCAGGCGAGTGTGGCGCCT 1220
QY 314 -----***HisThrGlyLeuGly***GlyIlePhe****IleLeuVal 329
Db 1221 GAAGCCCCACTGTGGGGGAGGGGTGGGGCT-----GTGCTTGCC 1262
QY 330 ThrGlyAlaVal-----AlaLeuAlaAlaTyrSerTyrPheArgIleAsnArgLysThr 347
Db 1263 GCTGAGCACTGTGCTTGGCTTGGTGGCGGAGCTCTTACCTCGTCCGCGGAGCAAGCCC 1322
QY 348 IleGlyPhe 350
Db 1323 ATGGGCTTT 1331

Search completed: April 1, 2004, 14:03:06
Job time : 1541 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 1, 2004, 09:56:50 ; Search time 121 Seconds
(without alignments)
1618.989 Million cell updates/sec

Title: US-09-466-778B-11

Perfect score: 1865

Sequence: 1 MTGPGHKCKSHYVDGL.....ALAAYSYFRINRKTIGFHP 353

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Xgapop 10.0, Xgapext 0.5
Xgapop 6.0, Xgapext 7.0
Delop 6.0, Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US09466778/runat_01042004_095643_10608/app_query.fasta_1.519
-DB=Issued_Patents_NA -QPMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=DILOSUM62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09466778 @CIGN 1 1 69 @runat_01042004_095643_10608 -NCFU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXI=7

Database : Issued Patents NA:
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	632.5	33.9	1482	4	US-09-016-434-230 Sequence 230, App
2	266.5	14.3	1414	1	US-08-024-868-1 Sequence 1, Appli
3	266.5	14.3	1414	2	US-08-242-097-1 Sequence 1, Appli
4	266.5	14.3	1414	3	US-09-206-695-1 Sequence 1, Appli
5	266.5	14.3	1414	4	US-09-799-118-1 Sequence 1, Appli
6	255.5	13.7	1734	4	US-09-484-970B-63 Sequence 63, Appl
7	200.5	10.8	1985	4	US-09-907-794A-212 Sequence 212, App
8	200.5	10.8	1985	4	US-09-905-125A-212 Sequence 212, App
9	200.5	10.8	1985	4	US-09-902-775A-212 Sequence 212, App
10	198.5	10.6	409	4	US-09-833-381-1252 Sequence 1252, App
11	157.5	8.4	1520	1	US-08-225-477B-1 Sequence 1, Appli
12	157.5	8.4	1520	5	PCT-US95-04353-1 Sequence 1, Appli

13	157.5	8.4	5191	1	US-08-340-428B-1 Sequence 1, Appli
14	157.5	8.4	5191	5	PCT-US93-07306-1 Sequence 1, Appli
15	155.5	8.3	3259	5	PCT-US95-03747-1 Sequence 1, Appli
16	155	8.3	1587	4	US-09-010-147B-19 Sequence 19, Appl
17	154.5	8.3	1400	2	US-08-001-078A-2 Sequence 2, Appli
18	154.5	8.3	1400	2	US-08-463-218-2 Sequence 2, Appli
19	154.5	8.3	1400	5	PCT-US94-00253-2 Sequence 2, Appli
20	153.5	8.2	1519	1	US-08-225-477B-2 Sequence 2, Appli
21	153.5	8.2	1519	5	PCT-US95-04353-2 Sequence 2, Appli
22	152.5	8.2	1720	4	US-09-148-545-53 Patent No. 5180808
23	150.5	8.1	8224	6	5180808-1 Sequence 111, App
24	122.5	6.6	1962	4	US-09-148-545-111 Sequence 28, Appl
25	122	6.5	1896	4	US-09-724-864-28 Sequence 13, Appl
26	109	5.8	2029	4	US-09-232-160-13 Sequence 200, App
27	109	5.8	2372	4	US-09-907-794A-200 Sequence 200, App
28	109	5.8	2372	4	US-09-905-125A-200 Sequence 200, App
29	109	5.8	2372	4	US-09-902-775A-200 Sequence 200, App
30	109	5.8	2404	4	US-09-833-381-849 Sequence 849, App
31	109	5.8	2404	4	US-09-833-381-853 Sequence 853, App
32	107	5.7	2313	2	US-08-892-880-1 Sequence 1, Appli
33	105	5.6	2430	2	US-08-820-170A-35 Sequence 35, Appl
34	105	5.6	2430	3	US-09-055-699-35 Sequence 35, Appl
35	105	5.6	2430	3	US-09-273-565-35 Sequence 35, Appl
36	105	5.6	2430	4	US-09-565-538-35 Sequence 35, Appl
37	105	5.6	2430	4	US-09-661-468-35 Sequence 35, Appl
38	105	5.6	2430	4	US-09-976-165-35 Sequence 35, Appl
39	105	5.6	2977	2	US-08-820-170A-36 Sequence 36, Appl
40	105	5.6	2977	3	US-09-055-699-36 Sequence 36, Appl
41	105	5.6	2977	3	US-09-273-565-36 Sequence 36, Appl
42	105	5.6	2977	4	US-09-565-538-36 Sequence 36, Appl
43	105	5.6	2977	4	US-09-661-468-36 Sequence 36, Appl
44	105	5.6	2977	4	US-09-976-165-36 Sequence 36, Appl
45	101.5	5.4	3207	1	US-07-946-497-1 Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-09-016-434-230
; Sequence 230, Application US/09016434
; Patent No. 6500938

; GENERAL INFORMATION:

; APPLICANT: Janice Au-Young

; APPLICANT: Jeffrey J. Sellhamer

; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING

; NUMBER OF SEQUENCES: 1490

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

; STREET: 3174 PORTER DRIVE

; CITY: PALO ALTO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/016,434

; FILING DATE: HEREWITH

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Zeller, Karen J.

; REGISTRATION NUMBER: 37,071

; REFERENCE/DOCKET NUMBER: PA-0002 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (650) 855-0555


```

; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 230:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1482 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: SPLNOT04
; CLONE: 1561663
US-09-016-434-230

Alignment Scores:
Pred. No.: 7,79e-73 Length: 1482
Score: 632.50 Matches: 143
Percent Similarity: 53.50% Conservative: 48
Best Local Similarity: 40.06% Mismatches: 145
Query Match: 33.91% Indels: 22
DB: 4 Gaps: 7

US-09-466-778B-11 (1-353) x US-09-016-434-230 (1-1482)

QY 8 LysCysGluCysLysSerHisTyrValGlyAspGlyLeuAsnCys---GluProGluGln 26
DB 4 CGCTGTGATGCGCAGCGAGGTACGTAGCGATGAGTGCAGTGTCTGAGGAGTCGGAA 63

QY 27 LeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHisAlaAspAlaLysCysVal 46
DB 64 CCACCTGTGGACCGCTGCTTGGCCAGCCAGCCGCTGCCACTCAGATGCCATGTGCCT 123

QY 47 AspleuHisPheGlnAspThrThrValGlyValPheHisLeuArgSerProLeuGlyGln 66
DB 124 GACCTGCATCTCCAGAGAAAGCGGCTGCGGCTTTTCCACCTCCAGCCAGCCAGCGCCCT 183

QY 67 TyrLysLeuThrPheAspAlaArgGluAlaCysAlaAsnGluAlaAlaThrMetAla 86
DB 184 TATGCTCTGAATTTTCGAGGCTCAGCGGCGATCGGA-AGACAGGGAGCGCTCTTCT 242

QY 87 ThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHisLeuCysSerAlaGlyTyr 106
DB 243 TCATTCCTCAGCTCTCTGCTGCCAGCAGCTGGGCTCCACCTGTGCTCATGGGCTGG 302

QY 107 LeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSerGlnAsnCysGlySerGly 126
DB 303 CTGGCCAAATGGCTCCACTGCCACCCCTGTGGTTTTCCTGTCGGCGAGCTGTGGCAATGGT 362

QY 127 ValValGlyIleValAspTyrGlyProArgProAsnLysSerGluMetTyrAspValPhe 146
DB 363 CGGTGGGCATAGTCAGCTGGTCCCGCAAGAACCTCTCAGAACGCTGGGATGCCTAC 422

QY 147 CysTyrArgMetLysAspValAsnCysThr***LysValGlyTyrValGlyAspGlyPhe 166
DB 423 TGCCTCCGTGTGCAAGATGTGGCTGCGGATGCCGAAATGGCTTCGCTGGGTGACGGGATC 482

QY 167 Ser---TyrSerGlyAsnLeuGlnValLeuMetSerPheProSerLeuThrAsnPhe 185
DB 483 AGCAGTCGAATGGGAAGCTGCTGATGTGCTGGCTGCCACTGCCCACTTCTCCACCTTC 542

QY 186 LeuThrGluValLeuAlaTyrSerAsnSerAlaArgGlyArgAlaPheLeuGluHis 205
DB 543 TATGGGATGCTATTGGGCTATGCCAATGCCACCCAGCGGGGTCTCGACTTCTTGACTTC 602

QY 206 LeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsnSerGlyLeuGlyGlu 225
DB 603 CTGGATGATGAGCTCACGTATAGACACTCTCTGCTCCCTGTCAATGAAGGCTTTTGGGAC 662

QY 226 AsnGluThrLeuSerGlyArgAspIleGluHisLeuAlaAsnValSerMetPhePhe 245
DB 663 AACATGAGCTGAGTGGCCAGCTTGGAGCTGCGCTCCAGCCAGCCCTCTCAAGT 722

QY 246 TyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGlySerLysLeuLeuile 265
DB 723 GCCAAC---GCCAGCCAGGGGAGTGTCTTCGGGCCCTCAGGCTCAGCTCATCATC 779

```

```

QY 266 ThrAspArg-----GlnAspProLeuHisProThrGluThrArgCysValAspGly 282
DB 780 AGTGACGACGCGCCTGACAAACAGTTCCTGGGCGCCTGTGGCCCCAGGAGTGTGGTT 839

QY 283 ArgAspThrLeuGluTyrAspIleCysAlaSerAsnGlyIleThrHisValIleSerArg 302
DB 840 AGCGGTATCATTTGTGGGACATCATGGCCTTCAATGGCATCATCCATGCTCTGGCCAGC 899

QY 303 ***LeuLysAlaProAlaProAlaProValThrLeu-----***His 315
DB 900 CCCCTCTCTGACACCCCCACAGCCCGGCGAGTGTGGCGCTGAAGCCCCACCTGTGGCG 959

QY 316 ThrGlyLeuGly***GlyIlePhe*****IleIleLeuValThrGlyAlaVal----- 333
DB 960 GCAGGCGTGGGGCT-----GTGCTTGGCGCTGGAGCACTGCTTGGC 1001

QY 334 AlaLeuAlaAlaTyrSerTyrPheArgIleAsnArgIleThrIleGlyPhe 350
DB 1002 TTGTGGCGGAGGCTCTCTACCTCCGTCGCGGAGGCAAGCCACGGGCTTT 1052

RESULT 2
US-08-024-868-1
; Sequence 1, Application US/08024868
; Patent No. 5386013
; GENERAL INFORMATION:
; APPLICANT: Lee, Tae Ho
; APPLICANT: Wisniewski, Hans Georg
; APPLICANT: Vilcek, Jan
; TITLE OF INVENTION: Cytokine-Induced Protein, TSG-6, DNA
; TITLE OF INVENTION: Coding Thereof and Uses Thereof
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESS: Browdy and Neimark
; CITY: Washington
; STATE: DC
; COUNTRY: US
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/024,868
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/642,312
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Livnat, Shmuel
; REGISTRATION NUMBER: 33,949
; REFERENCE/DOCKET NUMBER: VILCEK-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-628-5197
; TELEFAX: 212-737-3528
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1414 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; CELL TYPE: Fibroblast
; CELL LINE: FS-4
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 69..899
; OTHER INFORMATION:

```

US-08-024-868-1

Alignment Scores:

Pred. No.: 9,99e-25 Length: 1414
Score: 266.50 Matches: 64
Percent Similarity: 47.42% Conservative: 28
Best Local Similarity: 32.99% Mismatches: 78
Query Match: 14.29% Indels: 24
DB: 1 Gaps: 4

US-09-466-778B-11 (1-353) x US-08-024-868-1 (1-1414)

```
QY 52 AspThrThrValGlyValPheHisLeuArgSerProLeuGlyGlnTyrIleValLeuThrPhe 71
   :::::::::::::::::::::
Db 162 GAACGAGCAGCCGGTGTGTACCAAGAGCAGCGCTGCGCAATATCAACAGCTCACCTAC 221
   :::::::::::::::::::::
QY 72 AspValAlaArgGluAlaCysAlaAsnGluAlaAlaThrMetAlaThrTyrAsnGlnLeu 91
   :::::::::::::::::::::
Db 222 GCAGAGCTAAGCGGTGTGAATTTGAAGCGCGCCATCTCGCACTTACAAGCAGCTA 281
   :::::::::::::::::::::
QY 92 SerTyr***GlnIysAlaLysTyrHisLeuCysSerAlaGlyTyrLeuGluThrGlyArg 111
   :::::::::::::::::::::
Db 282 GAGCAGCCAGAAAAATTTGATTTTCATGCTGTGCTGCTGGATGGCTAAGGCGCAGA 341
   :::::::::::::::::::::
QY 112 ValAlaTyrProThrAlaPheAlaSerGlnAsnCysGlySerGlyValValGlyIleVal 131
   :::::::::::::::::::::
Db 342 GTTGGATACCCCATTTGTGAAGCCAGGCGCCCACTGTGGATTTGGAAAAAAGCTGCATTAT 401
   :::::::::::::::::::::
QY 132 AspTyrGlyProArgProAsnLysSerGluMetTyrAspValPheCysTyrArgMetLys 151
   :::::::::::::::::::::
Db 402 GATTATGGATCCGCTCTCAATAGAGTGAAGATGGATGGATGCTTACTTACTACACCCAC 461
   :::::::::::::::::::::
QY 152 AspValAsnCys-----Thr***IysValGlyTyrValGlyAspGlyPhe 166
   :::::::::::::::::::::
Db 462 GCAAGAGAGTGTGTGGCGCTTTTACAGATCCAAAGCGAATTTTAAATCTCCAGGCTTC 521
   :::::::::::::::::::::
QY 167 -----SerTyrSerGlyAsnLeuLeu----- 173
   :::::::::::::::::::::
Db 522 CCAATGAGTACCAAGATACCAAAATCTGCTACTGCGCACATTAAGTATGTCAG 581
   :::::::::::::::::::::
QY 174 GlnValLeuMetSerPheProSer-LeuThrAsnPheLeuThrGluVal----- 189
   :::::::::::::::::::::
Db 582 CGTATTCACCTGAGTTTGTAGATTTTGTACCTTGAAGATGACCCAGCTTGTGGCTGAT 641
   :::::::::::::::::::::
QY 190 -----LeuAlaTyrSerAsnSerSerAlaArgGlyArgAlaPheLeuGluHisLeuThrAs 208
   :::::::::::::::::::::
Db 642 TATGTTGAAATATATGACAGTTACGATGATGCCATGCTTGTGGGAAGATACCTGTGGA 701
   :::::::::::::::::::::
QY 208 pLeuSerIleArgGlyThrLeuPheValProGlnAsnSer 221
   :::::::::::::::::::::
Db 702 GATGAGCTTCCAGATGACATCATCAGTACGACGAGGAATGTCA 741
   :::::::::::::::::::::
```

RESULT 3

US-08-242-097-1
; Sequence 1, Application US/08242097
; Patent No. 5846763
; GENERAL INFORMATION:
; APPLICANT: Lee, Tae Ho
; APPLICANT: Wisniewski, Hans Georg
; APPLICANT: Vilcek, Jan
; TITLE OF INVENTION: Cytokine-Induced Protein, TSG-6, DNA Coding
; TITLE OF INVENTION: Therefor and Uses Thereof
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/242,097
FILING DATE: 13-MAY-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: LEE26/VILCEK=LB
TELEPHONE: 212-628-5197
TELEFAX: 212-737-3528

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1414 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
CELL TYPE: Fibroblast
CELL LINE: FS-4
FEATURE:

NAME/KEY: CDS
LOCATION: 69..899
OTHER INFORMATION:
US-08-242-097-1

Alignment Scores:

Pred. No.: 9,99e-25 Length: 1414
Score: 266.50 Matches: 64
Percent Similarity: 47.42% Conservative: 28
Best Local Similarity: 32.99% Mismatches: 78
Query Match: 14.29% Indels: 24
DB: 1 Gaps: 4

US-09-466-778B-11 (1-353) x US-08-242-097-1 (1-1414)

```
QY 52 AspThrThrValGlyValPheHisLeuArgSerProLeuGlyGlnTyrIleValLeuThrPhe 71
   :::::::::::::::::::::
Db 162 GAACGAGCAGCCGGTGTGTACCAAGAGCAGCGCTGCGCAATATCAACAGCTCACCTAC 221
   :::::::::::::::::::::
QY 72 AspValAlaArgGluAlaCysAlaAsnGluAlaAlaThrMetAlaThrTyrAsnGlnLeu 91
   :::::::::::::::::::::
Db 222 GCAGAGCTAAGCGGTGTGAATTTGAAGCGCGCCATCTCGCACTTACAAGCAGCTA 281
   :::::::::::::::::::::
QY 92 SerTyr***GlnIysAlaLysTyrHisLeuCysSerAlaGlyTyrLeuGluThrGlyArg 111
   :::::::::::::::::::::
Db 282 GAGCAGCCAGAAAAATTTGATTTTCATGCTGTGCTGCTGGATGGCTAAGGCGCAGA 341
   :::::::::::::::::::::
QY 112 ValAlaTyrProThrAlaPheAlaSerGlnAsnCysGlySerGlyValValGlyIleVal 131
   :::::::::::::::::::::
Db 342 GTTGGATACCCCATTTGTGAAGCCAGGCGCCCACTGTGGATTTGGAAAAAAGCTGCATTAT 401
   :::::::::::::::::::::
QY 132 AspTyrGlyProArgProAsnLysSerGluMetTyrAspValPheCysTyrArgMetLys 151
   :::::::::::::::::::::
Db 402 GATTATGGATCCGCTCTCAATAGAGTGAAGATGGATGGATGCTTACTTACTACACCCAC 461
   :::::::::::::::::::::
QY 152 AspValAsnCys-----Thr***IysValGlyTyrValGlyAspGlyPhe 166
   :::::::::::::::::::::
Db 462 GCAAGAGAGTGTGTGGCGCTTTTACAGATCCAAAGCGAATTTTAAATCTCCAGGCTTC 521
   :::::::::::::::::::::
QY 167 -----SerTyrSerGlyAsnLeuLeu----- 173
   :::::::::::::::::::::
Db 522 CCAATGAGTACCAAGATACCAAAATCTGCTACTGCGCACATTAAGTATGTCAG 581
   :::::::::::::::::::::
QY 174 GlnValLeuMetSerPheProSer-LeuThrAsnPheLeuThrGluVal----- 189
   :::::::::::::::::::::
Db 582 CGTATTCACCTGAGTTTGTAGATTTTGTACCTTGAAGATGACCCAGCTTGTGGCTGAT 641
   :::::::::::::::::::::
QY 190 -----LeuAlaTyrSerAsnSerSerAlaArgGlyArgAlaPheLeuGluHisLeuThrAs 208
   :::::::::::::::::::::
```

Db 642 TAAGTTGAAATATGACAGATTCACAGATGATGCCATGGCTTTGTGGGAAGATACTGTGGGA 701

Qy 208 pleuSerileArgGlyThrLeuPheValProGlnAsnSer 221

Db 702 GATGAGCTTCCAGATGACATCATCATGACAGAAATGTCA 741

RESULT 4

US-09-206-695-1

; Sequence 1, Application US/09206695

; Patent No. 6210905

; GENERAL INFORMATION:

; APPLICANT: Lee, Tae Ho

; APPLICANT: Wisniewski, Hans Georg

; APPLICANT: Vilcek, Jan

; TITLE OF INVENTION: Cytokine-Induced Protein, TSG-6, DNA Coding

; TITLE OF INVENTION: Thereof and Uses Thereof

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Browdy and Neimark

; STREET: 419 Seventh Street, N.W., Suite 300

; CITY: Washington

; STATE: D.C.

; ZIP: 20004

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/206,695

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/242,097

; FILING DATE: 13-MAY-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: BROWDY, Roger L.

; REGISTRATION NUMBER: 25,618

; REFERENCE/DOCKET NUMBER: LEE26/VILCEK=1B

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-628-5197

; TELEFAX: 212-737-3528

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1414 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; ORIGINAL SOURCE:

; ORGANISM: Homo sapiens

; CELL TYPE: Fibroblast

; CELL LINE: FS-4

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 69..899

; OTHER INFORMATION:

US-09-206-695-1

Alignment Scores:

Pred. No.:

Score: 9.99e-25

Length: 1414

Matches: 64

Conservatives: 28

Percent Similarity: 47.42%

Best Local Similarity: 32.99%

Mismatches: 78

Query Match: 14.29%

Indels: 24

Gaps: 4

DB:

US-09-466-778B-11 (1-353) x US-09-206-695-1 (1-1414)

Qy 52 AspThrValGlyValPheHisLeuArgSerProLeuGlyGlnTyrHisLeuThrPhe 71

Db 162 GAACGAGCAGCGGGTGTGTACACAGAGACGAGTCTGTGGCAATACAGCTCACCTAC 221

Qy 72 AspYsAlaArgGluAlaCysAlaAsnGluAlaAlaThrMetAlaThrTyrAsnGlnLeu 91

Db 222 GCAGAAGCTAAGCGCGGTGTGAATTTGAAGCGCGCATCTCGCAACTTACAGCAGCTA 281

Qy 92 SerTyr***GlnYsAlaLysTyrHisLeuCysSerAlaGlyTrpLeuGluThrGlyArg 111

Db 282 GAGCAGCAGCAAAATTTGGATTTTCATGTCTGTCTGATGGATGGCTTAAGGGCAGA 341

Qy 112 ValAlaTyrProThrAlaPheAlaSerGlnAsnCysGlySerGlyValValGlyVal 131

Db 342 GTTGATACCCCATTTGCAAGCCAGGCCCCCACTGTGGATTTGGAAAACTGGCATATT 401

Qy 132 AspTyrGlyProArgProAsnLysSerGluMetTrpAspValPheCysTyrArgMetLys 151

Db 402 GATTATGGAATCCGCTCTCATAGAGTCAAGATGGGATGCTATTGCTACACCCACAC 461

Qy 152 AspValAsnCys-----Thr***LysValGlyTyrValGlyAspGlyPhe 166

Db 462 CAAAGGAGTGTGTGGCGCTTTTACAGATCCAAAGCGAATTTTAAATCTCCAGGCTTC 521

Qy 167 -----SerTyrSerGlyAsnLeu----- 173

Db 522 CCAATGAGTACGAGATACCAATCTGCTACTGGCACATTAGACTCAAGTATGGTCAG 581

Qy 174 GlnValLeuMetSerPheProSer-LeuThrAsnPheLeuThrGluVal----- 189

Db 582 CGTATTCACCTGAGTCTTTTATGATTTTGCACCTTGAAGATGACCCAGGTTGCTGGCTGAT 641

Qy 190 ---LeuAlaTyrSerAsnSerSerAlaArgGlyArgAlaPheLeuGluHisLeuThrAs 208

Db 642 TATGTTGAATATATGACAGTACGATGATGCTGATGCTGCTTTGGGGAAGATACTGTGGA 701

Qy 208 pleuSerileArgGlyThrLeuPheValProGlnAsnSer 221

Db 702 GATGAGCTTCCAGATGACATCATCATGACAGAAATGTCA 741

RESULT 5

US-09-799-118-1

; Sequence 1, Application US/09799118

; Patent No. 6518401

; GENERAL INFORMATION:

; APPLICANT: Lee, Tae Ho

; Wisniewski, Hans Georg

; Vilcek, Jan

; TITLE OF INVENTION: Cytokine-Induced Protein, TSG-6, DNA Coding

; Thereof and Uses Thereof

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Browdy and Neimark

; STREET: 419 Seventh Street, N.W., Suite 300

; CITY: Washington

; STATE: D.C.

; ZIP: 20004

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/799,118

; FILING DATE: 06-Mar-2001

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/242,097

; FILING DATE: 13-MAY-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: BROWDY, Roger L.

; REGISTRATION NUMBER: 25,618

; REFERENCE/DOCKET NUMBER: LEE26/VILCEK=1B

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-628-5197

; TELEFAX: 212-737-3528

; INFORMATION FOR SEQ ID NO: 1:

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1414 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; CELL TYPE: Fibroblast
; CELL LINE: FS-4
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 69..899
; OTHER INFORMATION:
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-799-118-1

Alignment Scores:
Pred. No.: 9,99e-25 Length: 1414
Score: 266.50 Matches: 64
Percent Similarity: 47.42% Conservative: 28
Best Local Similarity: 32.99% Mismatches: 78
Query Match: 14.29% Indels: 24
DB: 4 Gaps: 4

US-09-466-778B-11 (1-353) x US-09-799-118-1 (1-1414)
Qy 52 AspThrThrValGlyValPheHisLeuArgSerProLeuGlyGlnTyrLysLeuThrPhe 71
Db 162 GAACGAGCAGCGGTGTGTACACAGAGAGCAGCGTCTGGCAATACAACTCACCTAC 221
Qy 72 AspLysAlaArgGluAlaCysAlaAsnGluAlaThrMetAlaThrTyrAsnGlnLeu 91
Db 222 GCAGAGCTAAGCGCGGTGTGCAATTTGAAGCGCGCCATCTCGCAACTTACAAGCAGCTA 281
Qy 92 SerTyr***GlnLysAlaLysTyrHisLeuCysSerAlaGlyTyrLeuGluThrGlyArg 111
Db 282 GAGCGAGCCAGAAAATTTGGATTTCATGCTGTGCTGTGGATGGATGGCTAAGGGCAGA 341
Qy 112 ValAlaTyrProThrAlaPheAlaSerGlnAsnCysGlySerGlyValValGlyLeVal 131
Db 342 GTTGGATACCCCATTTGTAAGCCAGGCGCCCACTGTGGATTGGAAAACTGGCATTATT 401
Qy 132 AspTyrGlyProArgProAsnLysSerGluMetTyrAspValPheCysTyrArgMetLys 151
Db 402 GATTATGGAATCCCTCTCAATAGAGTGAAGATGGGATGGCTATTGTACAACCCACAC 461
Qy 152 AspValAsnCys-----Thr***LysValGlyTyrValGlyAspGlyPhe 166
Db 462 GCAAGAGGTGTGTGGCGCTTTTACAGATCCAAAGCGAATTTTAAATCTCCAGGCTTC 521
Qy 167 -----SerTyrSerGlyAsnLeu----- 173
Db 522 CCAATAGTACGAAGATAACAAATCTGCTACTGGCAGATTAGACTCAAGTATGGTCAG 581
Qy 174 GlnValLeuMetSerPheProSer-LeuThrAsnPheLeuThrGluVal----- 189
Db 582 CGTATTCACCTGAGTATTTTAGATTGACCTTGAAGATGACCCAGGTGTGCTGGCTGAT 641
Qy 190 ----LeuAlaTyrSerAsnSerSerAlaArgGlyArgAlaPheLeuGluHisLeuThrAs 208
Db 642 TATGTTGAAATATATGACATTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 701
Qy 208 pLeuSerIleArgGlyThrLeuPheValProGlnAsnSer 221
Db 702 GATGAGCTTCCAGATGACATCATCAGTACAGGAAATGTCA 741

RESULT 6
US-09-484-970B-63
; Sequence 63, Application US/09484970B
; Patent No. 6426186
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen A.

; SEQUENCE CHARACTERISTICS:
; APPLICANT: Volkmut, Wayne
; APPLICANT: Walker, Michael G.
; TITLE OF INVENTION: BONE REMODELING GENES
; FILE REFERENCE: PB-0014 US
; CURRENT APPLICATION NUMBER: US/09/484,970B
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PERL Program
; SEQ ID NO 63
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6426186 201752.1
; NAME/KEY: unsure
; LOCATION: 1444, 1455, 1457, 1461, 1463, 1465, 1467, 1497, 1500, 1506, 1521, 1534-
; OTHER INFORMATION: a, c, g, or other
US-09-484-970B-63

Alignment Scores:
Pred. No.: 3.92e-23 Length: 1734
Score: 255.50 Matches: 64
Percent Similarity: 47.18% Conservative: 28
Best Local Similarity: 32.82% Mismatches: 78
Query Match: 13.70% Indels: 25
DB: 4 Gaps: 4

US-09-466-778B-11 (1-353) x US-09-484-970B-63 (1-1734)
Qy 52 AspThrThrValGlyValPheHisLeuArgSerProLeuGlyGlnTyrLysLeuThrPhe 71
Db 162 GAACGAGCAGCGGTGTGTACACAGAGAGCAGCGTCTGGCAATACAACTCACCTAC 221
Qy 72 AspLysAlaArgGluAlaCysAlaAsnGluAlaThrMetAlaThrTyrAsnGlnLeu 91
Db 222 GCAGAGCTAAGCGCGGTGTGCAATTTGAAGCGCGCCATCTCGCAACTTACAAGCAGCTA 281
Qy 92 SerTyr***GlnLysAlaLysTyrHisLeuCysSerAlaGlyTyrLeuGluThrGlyArg 111
Db 282 GAGCGAGCCAGAAAATTTGGATTTCATGCTGTGCTGTGGATGGATGGCTAAGGGCAGA 341
Qy 112 ValAlaTyrProThrAlaPheAlaSerGlnAsnCysGlySerGlyValValGlyLeVal 131
Db 342 GTTGGATACCCCATTTGTAAGCCAGGCGCCCACTGTGGATTGGAAAACTGGCATTATT 401
Qy 131 LAspTyrGlyProArgProAsnLysSerGluMetTyrAspValPheCysTyrArgMetLys 151
Db 402 TGATTATGGAATCCCTCTCAATAGAGTGAAGATGGGATGGCTATTGTGTACAAACCCACA 461
Qy 151 sAspValAsnCys-----Thr***LysValGlyTyrValGlyAspGlyPhe 166
Db 462 CGCAAGAGGTGTGTGGCGCTTTTACAGATCCAAAGCAAATTTTAAATCTCCAGGCTTC 521
Qy 166 e-----SerTyrSerGlyAsnLeu----- 173
Db 522 CCAATAGTACGAAGATAACAAATCTGCTACTGGCAGATTAGACTCAAGTATGGTCAG 581
Qy 174 GlnValLeuMetSerPheProSer-LeuThrAsnPheLeuThrGluVal----- 189
Db 582 CGGTATTCACCTGAGTATTTTAGATTGACCTTGAAGATGACCCAGGTGTGCTGGCTGTA 641
Qy 190 -----LeuAlaTyrSerAsnSerSerAlaArgGlyArgAlaPheLeuGluHisLeuThrA 208
Db 642 TTATGTTGAAATATATGACAGATTACGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 701
Qy 208 sLeuSerIleArgGlyThrLeuPheValProGlnAsnSer 221
Db 702 AGATGAGCTTCCAGATGACATCATCAGTACAGGAAATGTCA 742

RESULT 7
US-09-307-794A-212
; Sequence 212, Application US/09307794A
```

Patent No. 6635468
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Ashkenazi, Avi
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth, J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/907,794A
CURRENT FILING DATE: 2001-07-17
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 212
LENGTH: 1985
TYPE: DNA
ORGANISM: Homo sapiens
US-09-907-794A-212

Alignment Scores:
Pred. No.: 8,42e-16 Length: 1985
Score: 200.50 Matches: 84
Percent Similarity: 35.71% Conservative: 46
Best Local Similarity: 23.08% Mismatches: 93
Query Match: 10.75% Indels: 141
DB: 4 Gaps: 14
US-09-466-778B-11 (1-353) x US-09-907-794A-212 (1-1985)
QY 5 GlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeuAsnCysGluPro 24
DB 524 GGGCGTTACCGCTGTGAG-----GTCATTGACGGCGCTGGAGATGAA--- 565
QY 25 GluGlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHisAlaAspAlaLys 44
DB 566 -----AGCGGT 571
QY 45 CysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeuArgSerProLeu 64
DB 572 CTGGTGGAGCTGGAGCTGCGGGGTGTG-----GTCTTTCTCTTACCAAGTCCGCCAAC 622
QY 65 GlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsnGluAlaAlaThr 84
DB 623 GGGCGCTACCACTTCCAGGGCCGAGGAGTCTGTGAGCAGCAGCTGCGGTG 682
QY 85 MetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHisLeuCysSerAla 104
DB 683 GTGGCTCTCTTTGAGCAGCTCTCCGGGCGCTGGAGGAGGCGCTGGACTGGTGAACGCG 742
QY 105 GlyThrLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSerGlnAsnCysGly 124
DB 743 GGCTGGCTGCAGGATGCTACCGTCAGTACCCATCATGTTGCCCGCCGACCCCTGCGGT 802
QY 125 Ser---GlyValVal---GlyLeValAspTyrGlyProArgProAsnLysSerGluMet 142
DB 803 GGCCAGCGCTGGCACCTGGCGTGGGAAGTACGGCCCCCGCCACCGCGCGCTGCACGCG 862
QY 143 TrpAspValPheCysTyrArgMetLysAspValAsnCysThr***LysValGlyTyrVal 162
DB 863 TATGATGATTTCTGCTCGCT----- 883
QY 163 GlyAspGlyPheSerTyrSerGlyAsnLeuGlnValLeuMetSerPheProSerLeu 182
DB 883 ----- 883
QY 183 ThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArgAla--- 201
DB 884 -----ACTGCCCTCAAGGGCGGTGTAC 907
QY 202 PheLeuGluHisLeuThrAspLeuSerIleArg----- 212
DB 908 TACCTGGAGCACCTCGAGAGCTGACGCTGACAGAGGCAAGGAGGCGCTGCAGGAAGAT 967
QY 213 -----GlyThrLeuPheVal-ProGlnAsnSerGlyLeuGlyG1 225
DB 968 GATGCCAGCATCGCAAGTGGGAGAGCTCTTTCGGCTCGAAGTTCATGCGCTGGAC 1027
QY 225 uAsnGluThrLeuSerGlyArgAspIleGluHisLeuAlaAsnValSerMetPhePh 245
DB 1028 CGCTCGAGCGTGGCTGGCTGGCAGATG-----GCAGGCTCCGCTACCTGTGGTT 1078
QY 245 eTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGlySer-LysLeuLeu 265
DB 1079 CACCGCATCTCTAACTGTGGGC-----CCCGAGAGCGCTGGGTCCGAAGTTTGGC 1129
QY 265 leThrAspArgGlnAspProLeuHisProThrGluThrArgCysValAspGlyArgSept 285
DB 1130 TCCCCGACCCGACAGCGCGCTT-----GTACGGTGT----- 1161
QY 285 hrLeuGluTrpAspIleCysAlaSerAsnGlyIleThrHisValIleSerArg***LeuL 305
DB 1162 -----TTAC 1165


```
QY 225 uAsnGluThrLeuSerGlyArgAspIleGluHisLeuAlaAsnValSerMetPhePh 245
Db 1028 CGTGCAGCGTGGCTGCTGCAGATG-----GCAGGCTCCGTACCTGTGGTT 1078
QY 245 eTyraAsnAspLeuValAenGlyThrThrLeuGlnThrArgLeuGlySer-LysLeuLeu 265
Db 1079 CACCCGATCTTAAGTGTGGG-----CCCCAGAGCTGGGGTCCGAGCTTTGGC 1129
QY 265 leThrAspArgGlnAspProLeuHisProThrGluThrArgCysValAspGlyArgAsp 285
Db 1130 TTCCCGCAGCCGACAGCGCTT-----GTACGGTGT-----1161
QY 285 hrLeuGluTrpAspIleCysAlaSerAsnGlyIleThrHisValIleSerArg***Leu 305
Db 1162 -----TTAC 1165
QY 305 ysAlaProAlaProValThrLeu***HisThrGlyLeuGly***GlyIlePhe----- 323
Db 1166 TGTACCCGACCA-----CTAGGACCTGGGGCCCTCCCT 1201
QY 324 -----*****lelleLeuValThrGlyAlaValAlaLeuAlaIleYrs 339
Db 1202 GCGCATTCCTCACTGCTGTGTATTATTAGAGTGTTCGTTTCCCTTGTGGTTGGA 1261
QY 339 erTyrrPhe 341
Db 1262 GCCATTTT 1269

RESULT 9
US-09-902-775A-212
; Sequence 212, Application US/0902775A
; Patent No. 6686451
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary B.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/902,775A
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
```

```
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 212
; LENGTH: 1985
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-902-775A-212

Alignment Scores:
Pred. No.: 8,42e-16 Length: 1985
Score: 200.50 Matches: 84
Percent Similarity: 35.71% Conservative: 46
Best Local Similarity: 23.08% Mismatches: 93
Query Match: 10.75% Indels: 141
DB: 4 Gaps: 14

US-09-466-778B-11 (1-353) x US-09-902-775A-212 (1-1985)
QY 5 GlnLysHisLysCysGluCysLysSerHisTyrrValGlyAspGlyLeuAsnCysGluPro 24
Db 524 GGGCTTACCGCTGTGAG-----GTCTTTCCTTACCGTCCCAAC 565
QY 25 GlnGlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHisAlaAspAlaLys 44
Db 566 -----AGCGGT 571
QY 45 CysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeuArgSerProLeu 64
Db 572 CTGCTGAGCTGGAGCTCGGGGTGTG-----GTCTTTCCTTACCGTCCCAAC 622
QY 65 GlyGlnTyrrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsnGluAlaAlaThr 84
Db 623 GGGCGCTACCACTTCCAGCGGCGCAGAGCTCTGTGACAGAGGCTGCGGTG 682
QY 85 MetaLThrTyrrAsnGlnLeuSerTyrr***GlnLysAlaLysTyrrHisLeuCysSerAla 104
Db 683 GTGGCTTCCCTTTGAGCAGCTCTTCGGGGCTGGAGGAGGCGCTGGACTGCTGCAACGCG 742
QY 105 GlyTrpLeuGluThrGlyArgValAlaTyrrProThrAlaPheAlaSerGlnAsnCysGly 124
Db 743 GGCTGGCTGAGGATGCTACGTCAGTACCCATCATGTGCCCCGGCAGCCCTGCGGT 802
QY 125 Ser---GlyValVal---GlyIleValAspTyrrGlyProArgProAsnLysSerGluMet 142
Db 803 GGCCAGGCTGGCACCTGGCGTGGCGAGAGTACGGCCCCCGCCACCCCGCTGTCACCGC 862
QY 143 TrpAspValPheCysTyrrArgMetLysAspValAsnCysThr***LysValGlyTyrrVal 162
Db 863 TATCATGTATTCTGCTTCGCT-----883
```

QY 163 GlyAspGlyPheSerTyrSerGlyAsnLeuLeuGlnValLeuMetSerPheProSerLeu 182
 Db 883 ----- 883
 QY 183 ThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArgAla--- 201
 Db 884 -----ACTGCCCTCAAGGGCGGGTGTAC 907
 QY 202 PheLeuGluHisLeuThrAspLeuSerIleArg----- 212
 Db 908 TACCTGGAGCACCCTGAGAGAGCTGACCTGACAGAGGCAAGAGGAGGCTGCGAGAGAGAT 967
 QY 213 -----GlyThrLeuPheVal-ProGlnAsnSerGlyLeuGlyG1 225
 Db 968 GATGCCACGATCGCAAGGTGGGACAGCTCTTCCCGCTCGGAAGTTCCATGCGCTGGAC 1027
 QY 225 uAsnGluThrLeuSerGlyArgAspIleGluHisLeuAlaAsnValSerMetPhePh 245
 Db 1028 CGCTGCCAGCTGGCTGGCTGGCAGATG-----GCAGCGTCCGCTACCCCTGTGTT 1078
 QY 245 eTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGlySer-LysLeuLeuI 265
 Db 1079 CACCGCATCTACTCTGGGC-----CCCCAGAGCTGGGTCCGAGAGTTTGGC 1129
 QY 265 leThrAspArgGlnAspProLeuHisProThrGluThrArgCysValAspGlyArgAspT 285
 Db 1130 ITCCCGACCGCGAGCCGCTT-----GTACGGTGT----- 1161
 QY 285 hrLeuGluTfPAspIleCysAlaSerAsnGlyIleThrHisValIleSerArg***LeuL 305
 Db 1162 -----TTAC 1165
 QY 305 yslalProProAlaProValThrLeu***HisThrGlyLeuGly***GlyIlePhe---- 323
 Db 1166 TGCTACCGCAGCA-----CTAGGACCTGGGGCCCTCCCT 1201
 QY 324 -----*****IleIleLeuValThrGlyAlaValAlaLeuAlaAlaTyrS 339
 Db 1202 GCGCATTCCTCACTGGTGTATTATTAGTGTGCTTCCCTTTCCCTTTGGTGGT 1261
 QY 339 erTyrPhe 341
 Db 1262 GCCATTIT 1269
 RESULT 10
 US-09-833-381-1252
 ; Sequence 1252, Application US/09833381
 ; Patent No. 6672186
 ; GENERAL INFORMATION:
 ; APPLICANT: Robison, Keith E
 ; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
 ; FILE REFERENCE: 5800-119
 ; CURRENT APPLICATION NUMBER: US/09/833,381
 ; CURRENT FILING DATE: 2001-04-11
 ; PRIOR APPLICATION NUMBER: 09/516,448
 ; PRIOR FILING DATE: 2000-02-29
 ; NUMBER OF SEQ ID NOS: 2050
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 1252
 ; LENGTH: 409
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)...(409)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-833-381-1252

Alignment Scores:
 Pred. No.: 1.12e-16 Length: 409
 Score: 198.50 Matches: 35
 Percent Similarity: 64.18% Conservative: 8

Best Local Similarity: 52.24% Mismatches: 23
 Query Match: 10.64% Indels: 1
 DB: 4 Gaps: 1
 US-09-466-778B-11 (1-353) x US-09-833-381-1252 (1-409)
 QY 2 ThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeuAsn 21
 Db 209 ACCGGTCCGAAACACACGGGCTGTGAATGCCAGTAGGCTACGTGGGTGATGGCTGCAG 268
 QY 22 CysGluProGlu---GlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis 40
 Db 269 TGCTANAGAGAGTTGAACCCCTGTGGACAGATGTTGGAGAGATCATCCTTGCCAC 328
 QY 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60
 Db 329 ACTGATGCTTGTGCACTGACCTACCTATTTCCAGGAAAACAGGCTGGTGTCTTCCACATC 388
 QY 61 ArgSerProLeuGlyGlnTyr 67
 Db 389 CAGGCCACCACTGGCCCTTAT 409
 RESULT 11
 US-08-225-477B-1
 ; Sequence 1, Application US/08225477B
 ; Patent No. 5633370
 ; GENERAL INFORMATION:
 ; APPLICANT: Susan Hockfield
 ; APPLICANT: Diane M. Jaworski
 ; TITLE OF INVENTION: BEHAV, A Brain Hva-
 ; TITLE OF INVENTION: luronan-Binding Protein
 ; NUMBER OF SEQUENCES: 11
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: St. Onge Steward Johnston & Reens
 ; STREET: 986 Bedford Street
 ; CITY: Stamford
 ; STATE: CT
 ; COUNTRY: United States
 ; ZIP: 06905
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5" 1.44 Mb diskette
 ; COMPUTER: IBM PC
 ; OPERATING SYSTEM: MS DOS
 ; SOFTWARE: Word Processor
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/225,477B
 ; FILING DATE: April 8, 1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Mary M. Krinsky
 ; REGISTRATION NUMBER: 32423
 ; REFERENCE/DOCKET NUMBER: 1751-P0004
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 203-324-6155
 ; TELEFAX: 203-327-1096
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1520 bases
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE:
 ; DESCRIPTION: DNA encoding a protein
 ; FRAGMENT TYPE: entire sequence
 ; IMMEDIATE SOURCE: rat brain
 ; FEATURE:
 ; NAME/KEY: rat BEHAV
 US-08-225-477B-1

Alignment Scores:
 Pred. No.: 2.45e-10 Length: 1520
 Score: 157.50 Matches: 46
 Percent Similarity: 41.21% Conservative: 22
 Best Local Similarity: 27.88% Mismatches: 60

Query Match:	8.45%	Indels:	37
DB:	1	Gaps:	5
US-09-466-77BB-11 (1-353) x US-08-225-477B-1 (1-1520)			
QY	10	GlucylsersHisTyrValGlyAspGlyLeuAsnCysGluProGluGlnLeuProIle	29
Db	851	CAAGTGTATGGCTGGCTGCTCGACCAACCGTAGGTACCCATCCAGAACCCACGA	910
QY	30	AspArgCysLeuGlnAsp	35
Db	911	GAAGCGTGTATGGAGACATGGATGGTACCTCGAGTGGGAATTACGGAGTGGTGGGT	970
QY	36	-----AsnGlyGlnCysHisAlaAspAlaLysCysValAspLeuHisPhe	50
Db	971	CCTGATGATCTCTACGATGCTACTGTTATGCCGAA-----GACCTAAAT---	1015
QY	51	GlnAspThrThrValGlyValPheHisLeuArgSerProLeuGlyGlnTyrLysLeuThr	70
Db	1016	-----GGAGAACTGTTCTTAGTGGCCCTCCGSG-----AAGCTGACG	1054
QY	71	PheAspLysAlaArgGluAlaCysAlaAsnGluAlaThrMetAlaThrTyrAsnGln	90
Db	1055	TGGAGGAGGCTCGGACTACTGTCTGGAACCGCGTCTCAGATCGCTAGCAGCGGCCAG	1114
QY	91	LeuSerTyr***GlnLysAlaLysTyrHisLeuCysSerAlaGlyTrpLeuGluThrGly	110
Db	1115	CTATACGCGCATGGAAATGGCGCTTGGACAGATGTAGCCCTGGCTGGCTGGTATGGC	1174
QY	111	ArgValAlaTyrProThrAlaPheAlaSerGlnAsnCysGlySerGlyValValGlyIle	130
Db	1175	AGTGTGGGTACCCCATCATCAGCCACGACGCTGTGGGGAGAGCCCTGCCNAGGATC	1234
QY	131	ValAspTyrGlyProArgProAsn-----LysSerGluMetTrpAsp	144
Db	1235	AAGACCCTCTTCCTCTTCCCAACCAACAGACTGGCTTCCCAAGCAGACAGACCGCTTCAAT	1294
QY	145	ValPheCysTyrArg	149
Db	1295	GTCTACTGCTTCCGA	1309

RESULT 12
PCT-US95-04353-1
Sequence 1, Application PC/TUS9504353
GENERAL INFORMATION:
APPLICANT: Susan Hockfield
APPLICANT: Diane M. Jaworski
TITLE OF INVENTION: BEHAB, A Brain Hyaluronan-Binding Protein
TITLE OF INVENTION: luronan-Binding Protein
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: St. Onge Steward Johnston & Reems
STREET: 986 Bedford Street
CITY: Stamford
STATE: CT
COUNTRY: United States
ZIP: 06905
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" 1.44 Mb diskette
COMPUTER: IBM PC
OPERATING SYSTEM: MS DOS
SOFTWARE: Word Processor
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04353
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/225,477
FILING DATE: April 8, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Mary M. Krinsky
REGISTRATION NUMBER: 32423
REFERENCE/DOCKET NUMBER: 1751-P0004
TELECOMMUNICATION INFORMATION:

```

; TELEPHONE: 203-324-6155
; TELEFAX: 203-327-1096
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1520 bases
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: DNA encoding a protein
; FRAGMENT TYPE: entire sequence
; IMMEDIATE SOURCE: rat brain
; FEATURE:
; NAME/KEY: rat BEHAV
PCT-US95-04353-1

Alignment Scores:
Pred. No.: 2,45e-10 Length: 1520
Score: 157,50 Matches: 46
Percent Similarity: 41,21% Conservative: 22
Best Local Similarity: 27,88% Mismatches: 60
Query Match: 8,45% Indels: 37
DB: 5 Gaps: 5

US-09-466-778B-11 (1-353) x PCT-US95-04353-1 (1-1520)

Qy 10 GluCysLysSerHisTyrValGlyAspGlyLeuAsnCysGluProGluGlnLeuProIle 29
Db 851 CAGTGTGATCTGGCTGGTCTCGACCAAAACCTGAGTACCCCATCCAGAACCCACGA 910
Qy 30 AspArgCysLeuGlnAsp----- 35
Db 911 GAAGCTGTTATGGAGACATGGATGGCTACCTGGATGCGGAATTACGGAGTGGTGGT 970
Qy 36 -----AenGlyGlnCysHisAlaAspAlaLysCysValAspLeuHisPhe 50
Db 971 CCTGATGATCTACGATGTCCTACTGTTATGCCGAA-----GACCTAAAT--- 1015
Qy 51 GlnAspThrThrValGlyValPheHisLeuArgSerProLeuGlyGlnTyrLysLeuThr 70
Db 1016 -----GGAGAACTGTCCTAGTGCCTCCCTCCCGC-----AAGCTGACG 1054
Qy 71 PheAspLysAlaArgGluAlaCysAlaAsnGluAlaAlaThrMetAlaThrTyrAsnGln 90
Db 1055 TGGGAGAGCTCGGACTACTGCTGGAACGGCGTCTCAGATCGCTAGCAGCGGCAG 1114
Qy 91 LeuSerTyr**GlnLysAlaLysTyrHisLeuCysSerAlaGlyTrpLeuGluThrGly 110
Db 1115 CTATACCGCGCATGGAATGGCGGCTGGACAGATGTAGCCCTGCCTGGCTGATGGC 1174
Qy 111 ArgValAlaTyrProThrAlaPheAlaSerGlnAsnCysGlySerGlyValValGlyIle 130
Db 1175 AGTGTGGGTACCCCATCATCACGCCAGCAACGGCTGTGGGGAGGCCCTCCAGGAGTC 1234
Qy 131 ValAspTyrGlyProArgProAsn-----LysSerGluMetTrpAsp 144
Db 1235 AAGACCTCTTCCTCTTTTCCCAACCAAGACTGGCTTCCCAGCAAGACACCGCTCAAT 1294
Qy 145 ValPheCysTyrArg 149
Db 1295 GTCTACTGCTTCCGA 1309

RESULT 13
US-08-340-428B-1
; Sequence 1, Application US/08340428B
; Patent No. 5648465
; GENERAL INFORMATION:
; APPLICANT: MARGOLIS, Richard U.
; APPLICANT: RAUCH, Uwe
; APPLICANT: MARGOLIS, Renee K.
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES FOR A
; TITLE OF INVENTION: NEUROCAN AS A CHONDROITIN SULFATE PROTEOGLYCAN
; NUMBER OF SEQUENCES: 49

```


QY 17 GlyAspGlyLeuAsnGlyProGluGlnLeuProIleAspArgCys----- 32
Db TGGGGC-----CACTGCCAGCAGTCTCTGGG-CTGTATCGCTGCCAAGTGTGAAG 505
QY 33 ---LeuGlnAspAsnGlyGlnCysHisAlaAspAlaLysCysValAspLeuHisPheGln 51
Db :-----CAGGACCTGTGAACCTG 538
QY 52 AspThrThrValGlyValPheHisLeuArgSerProLeuGlyGlnTyrLysLeuThrPhe 71
Db :-----CACTGCCAGCAGTCTCTGGG-CTGTATCGCTGCCAAGTGTGAAG 505
QY 539 GAAGTACAGCGGCTCTGTCTTCCATTATCGCGGCGCGGACCGCTATCGCTGACCTTC 598
QY 72 AspLysAlaArgGluAlaCysAlaAsnGluAlaAlaThrMetAlaThrTyrAsnGlnLeu 91
Db :-----CAGGACCTGTGAACCTG 538
QY 599 GCTGAGCGCCAGAGGCTGTCACTGAGTCTCGCTACCATTCGCGCTCCAGGACCTG 658
QY 92 SerTyr**GlnLysAlaLysTyrHisLeuCysSerAlaGlyTyrLeuGluThrGlyArg 111
Db :-----CACTGCCAGCAGTCTCTGGG-CTGTATCGCTGCCAAGTGTGAAG 505
QY 659 CAGGCTGCGTTCGAGATGCTTTGACAACTGGATGCGGCTGCTCTCAGACCGCAG 718
QY 112 ValAlaTyrProThrAlaPheAlaSerGlnAsnCys-----GlySerGlyValVal 128
Db :-----CACTGCCAGCAGTCTCTGGG-CTGTATCGCTGCCAAGTGTGAAG 505
QY 719 GTCCCGTACCCGATCACTCAGTCCGCTCCGCTGCTATGCTGATCGCAGCAGCTGCCA 778
QY 129 GlyIleValAspTyrGlyProArgProAsnLysSerGluMetTyrAspValPheCysTyr 148
Db :-----CACTGCCAGCAGTCTCTGGG-CTGTATCGCTGCCAAGTGTGAAG 505
QY 779 GTGTCCGAGTACCGG---AGACGCGACCGCAGGAACTTACGATGCTTACGCTTT 835
QY 149 ArgMetLysAspValAsnCysThr**LysValGlyTyrValGly 163
Db :-----CTAGGGGCTGAAGTCTTTTACGTGGC 871

RESULT 15

PCT-US95-03747-1

; Sequence 1, Application PC/TUS9503747

; GENERAL INFORMATION:

; APPLICANT: LA JOLLA CANCER RESEARCH FOUNDATION

; TITLE OF INVENTION: Brevican, A Glial Cell Proteoglycan

; NUMBER OF SEQUENCES: 3

; CORRESPONDENCE ADDRESS:

; ADDRESS: Campbell and Flores

; STREET: 4370 La Jolla Village Drive, Suite 700

; CITY: San Diego

; STATE: California

; COUNTRY: USA

; ZIP: 92122

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US95/03747

; FILING DATE: 27-MAR-1995

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Imbra, Richard J.

; REGISTRATION NUMBER: 37,643

; REFERENCE/DOCKET NUMBER: FP-LJ 1453

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (619) 535-9001

; TELEFAX: (619) 535-8949

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3259 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 112...2848

; PCT-US95-03747-1

Alignment Scores:

Pred. No.: 1.59e-09 Length: 3259
Score: 155.50 Matches: 45
Percent Similarity: 42.77% Conservative: 23
Best Local Similarity: 28.30% Mismatches: 66
Query Match: 8.34% Indels: 25
DB: 5 Gaps: 5

US-09-466-778B-11 (1-353) x PCT-US95-03747-1 (1-3259)

QY 10 GluCysLysSerHisTyrValGlyAspGlyLeuAsnCysGluProGluGlnLeuProIle 29
Db :-----CACTGCCAGCAGTCTCTGGG-CTGTATCGCTGCCAAGTGTGAAG 505
QY 715 CAGTGTGACGCTGCTGTCCGACAGACCGTGAAGTATCCATCCAGAGCCAGA 774
Db :-----CACTGCCAGCAGTCTCTGGG-CTGTATCGCTGCCAAGTGTGAAG 505
QY 30 AspArgCys-----LeuGlnAspAsnGlyGlnCysHis 40
Db :-----CACTGCCAGCAGTCTCTGGG-CTGTATCGCTGCCAAGTGTGAAG 505
QY 775 GAGGCTCTGTATGGAGACATGATGCTTCCCTGGGGTCCGGAACCTACGAGTGTGTCGAC 834
Db :-----CACTGCCAGCAGTCTCTGGG-CTGTATCGCTGCCAAGTGTGAAG 505
QY 41 AlaAspAlaLysCysValAspLeuHis-----PheGlnAspThrThrValGly 56
Db :-----CACTGCCAGCAGTCTCTGGG-CTGTATCGCTGCCAAGTGTGAAG 505
QY 835 CCCGAT-----GACCTCTATGATGTTTACTGTTATGCTGAAGAATAAATGGA 882
Db :-----CACTGCCAGCAGTCTCTGGG-CTGTATCGCTGCCAAGTGTGAAG 505
QY 57 ValPheHisLeuArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGlu 76
Db :-----CACTGCCAGCAGTCTCTGGG-CTGTATCGCTGCCAAGTGTGAAG 505
QY 883 GAGCTGTTCCTGGTGGTCCCT-----CCAGACAAGCTGACCTTGGAGAGCGCGGACA 936
Db :-----CACTGCCAGCAGTCTCTGGG-CTGTATCGCTGCCAAGTGTGAAG 505
QY 77 AlaCysAlaAsnGluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLys 96
Db :-----CACTGCCAGCAGTCTCTGGG-CTGTATCGCTGCCAAGTGTGAAG 505
QY 937 TACTGCCAGAGCGGGTGTAAAGATTGCAACCCGCGGAGCTGTATGCGAGCTGGGAT 996
Db :-----CACTGCCAGCAGTCTCTGGG-CTGTATCGCTGCCAAGTGTGAAG 505
QY 97 AlaLysTyrHisLeuCysSerAlaGlyTyrLeuGluThrGlyArgValAlaTyrProThr 116
Db :-----CACTGCCAGCAGTCTCTGGG-CTGTATCGCTGCCAAGTGTGAAG 505
QY 997 GGTGCGCTGGACCGCTGCGAGCTCTGGCTGGCTGTCTGATGGCAGTGTGCGCTACCCCATC 1056
Db :-----CACTGCCAGCAGTCTCTGGG-CTGTATCGCTGCCAAGTGTGAAG 505
QY 117 AlaPheAlaSerGlnAsnCysGlySerGlyValValGlyIle----- 130
Db :-----CACTGCCAGCAGTCTCTGGG-CTGTATCGCTGCCAAGTGTGAAG 505
QY 1057 GTCACCCCGCAGCGCGCTGTGGTGGGGCGCTCCCTGGTGTCAAGACTCTCTCTCTCTTC 1116
Db :-----CACTGCCAGCAGTCTCTGGG-CTGTATCGCTGCCAAGTGTGAAG 505
QY 131 ValAspTyrGlyProArgProAsnLysSerGluMetTyrAspValPheCysTyrArg 149
Db :-----CACTGCCAGCAGTCTCTGGG-CTGTATCGCTGCCAAGTGTGAAG 505
QY 1117 CCCAACCCAGACTGGCTTCCCAACCAAGCAGCAGCGCTTCAACGCTTACTGCTTCCGA 1173

Search completed: April 1, 2004, 09:59:16

Job time : 132 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 1, 2004, 09:56:50 ; Search time 549 Seconds
(without alignments)

2731.537 Million cell updates/sec

Title: US-09-466-778B-11

Perfect score: 1865

Sequence: 1 MTGPGKHCKECKSHYVDGL.....ALAAYSYFRINRKTIGXHF 353

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US09466778/runat_01042004_095641_10565/app_query.fasta_1.519
-DB=N Geneseq 29Jan04 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi
-LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -NCPU=6 -ICPU=3
-USER=US09466778 @CNG 1.1.470 @runat_01042004_095641_10565 -NCPUL=0
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_29Jan04.*

1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002s.*
7: Geneseqn2003as.*
8: Geneseqn2003bs.*
9: Geneseqn2003cs.*
10: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1849	99.1	1259	3	AAA57365
2	1772	95.0	4576	6	ABA04662
3	1772	95.0	4642	7	ABX13822
4	1772	95.0	4962	7	ABX13823
5	1769	94.9	3260	6	ABT08489
6	1558	83.5	3625	4	AAF87120
7	1558	83.5	8495	6	ABT08488
8	1554	83.3	8444	6	ABT08487

9	1514.5	81.2	4706	6	ABA04648	AbA04648 Rat Hyalu
10	1514.5	81.2	4706	7	ABX13821	Abx13821 cDNA enco
11	1275	68.4	1377	3	AAC76373	Aac76373 Human ORF
12	1203	64.5	2011	4	AAF87113	Aaf87113 NOV2 codi
13	661.5	35.5	3681	6	ABL90758	AbL90758 Human poi
14	658.5	35.3	7879	8	ACF04000	Acf04000 Human CLE
15	656.5	35.2	1522	3	AAA57363	Aaa57363 DNA encod
16	645.5	34.6	2753	3	AAA57362	Aaa57362 DNA encod
17	645.5	34.6	2753	9	ADB47730	Adb47730 Novel hum
18	642	34.4	2160	9	ADB63670	Adb63670 Human CDN
19	632.5	33.9	1482	7	ACA55632	Aca55632 Human sig
20	626	33.6	1804	4	AAF87114	Aaf87114 NOV3 codi
21	581	31.2	2863	9	ADB62456	Adb62456 Human CDN
22	535	28.7	2483	4	AAF87119	Aaf87119 NOV8 codi
23	535	28.7	2512	4	AAF87117	Aaf87117 NOV6 codi
24	363	19.5	474	8	ACH45565	Ach45565 Human foe
25	272.5	14.6	781	7	ABX74436	Abx74436 Human CDN
26	267.5	14.3	1144	7	ABX63446	Abx63446 Human CDN
27	267.5	14.3	1144	9	ADE25720	Ade25720 Human CDN
28	267.5	14.3	1422	7	ACC72678	Acc72678 Human can
29	267.5	14.3	1422	7	ABX76382	Abx76382 Lung canc
30	267.5	14.3	1728	7	ABX63445	Abx63445 Human CDN
31	266.5	14.3	1414	2	AAQ27190	Aaq27190 Tumour ne
32	266.5	14.3	1414	2	AAV71778	Aav71778 Tumour ne
33	266.5	14.3	1414	4	AAD06019	Aad06019 Human tum
34	266.5	14.3	1414	6	ABSS4635	Abss4635 Human CDN
35	256.5	13.8	1411	4	AAH23114	Aah23114 Osteoarth
36	255.5	13.7	1734	6	ABST70406	Abst70406 Human bon
37	248	13.3	351	6	ABV95472	Abv95472 Human pan
38	243.5	13.1	1414	6	ABL68341	AbL68341 Kidney ca
39	243.5	13.1	1414	6	ABK83989	Abk83989 Human CDN
40	243.5	13.1	1414	7	ACC72679	Acc72679 Human can
41	243.5	13.1	1414	7	ABX76383	Abx76383 Lung canc
42	231	12.4	396	8	ACH30555	Ach30555 Human tes
43	213	11.4	718	5	AAS00854	Aas00854 Human CDN
44	202.5	10.9	1365	4	AAI58121	Aai58121 Human pol
45	200.5	10.8	1984	2	AAK52252	Aak52252 Protein p

ALIGNMENTS

RESULT 1
AAK57365
ID AAA57365 standard; DNA; 1259 BP.

XX
AC AAA57365;

DT 03-OCT-2000 (first entry)

DE DNA encoding a human hyaluronan-binding protein, designated BM-HABP.

XX
KW Hyaluronan-binding protein WF-HABP; OE-HABP; BM-HABP;
KW proliferative condition; metastasis; inflammation; ischemia;
KW host defence dysfunction; immune surveillance dysfunction; arthritis;
KW multiple sclerosis; autoimmunity; immune dysfunction; allergy; ss.

XX
OS Homo sapiens.

XX
PH Key

FT CDS

FT
FT Location/Qualifiers
FT 199..1257
FT /tag= a
FT /product= "hyaluronan-binding protein"
FT /transl_except= (pos: 478..480, aa: Xaa)
FT /transl_except= (pos: 667..669, aa: Xaa)
FT /transl_except= (pos: 1105..1107, aa: Xaa)
FT /transl_except= (pos: 1138..1140, aa: Xaa)
FT /transl_except= (pos: 1158..1159, aa: Xaa)
FT /transl_except= (pos: 1168..1170, aa: Xaa)
FT /transl_except= (pos: 1171..1173, aa: Xaa)
FT /transl_except= (pos: 1249..1251, aa: Xaa)
FT /note= "Xaa is an unspecified amino acid; no termination
codon given"

PN W0200039166-AL.
 XX PD 06-JUL-2000.
 XX PF 20-DEC-1999; 99WO-US030462.
 XX PR 23-DEC-1998; 98US-0113871P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX (AMNA-) AMERICAN NAT RED CROSS.
 XX Hastings GA, Liau G, Tsifrina E;
 XX WPI; 2000-452376/39.
 XX P-PSDB; AAY93913.
 XX New hyaluron-binding proteins, known as full-length WF-HABP, WF-HABP, OE-HABP, and BM-HABP, useful for treating proliferative conditions, metastasis, inflammation, ischemia, arthritis and multiple sclerosis.
 XX Claim 2; Fig 4A-B; 457pp; English.
 XX The present sequence encodes a hyaluronan-binding protein. The specification describes four hyaluronan-binding proteins, known as WF-HABP, WF-HABP, OE-HABP, and BM-HABP. The polypeptides are useful for treating diseases such as proliferative conditions, metastasis, inflammation, ischemia, host defence dysfunction, immune surveillance, arthritis, multiple sclerosis, autoimmunity, immune dysfunction and allergy
 XX SQ Sequence 1259 BP; 295 A; 348 C; 346 G; 258 T; 0 U; 12 Other;
 Alignment Scores:
 Pred. No.: 5,24e-203 Length: 1259
 Score: 1849.00 Matches: 353
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 99.14% Indels: 0
 DB: 3 Gaps: 0
 US-09-466-778B-11 (1-353) x AAAS7365 (1-1259)
 QY 1 MetThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeu 20
 DB 199 ATGACAGCCCGGCGACACAGAGTGTGAGTGAAGTCACTATGTGGAGATGGGCTG 258
 QY 21 AsnCysGluProGluGlnLeuProIleAspArgCysLeuGlnAspAsnGlyClnCysHis 40
 DB 259 AACTGTGAGCCCGGAGCAGCTGCCATTGACCGCTGCTTACAGGACATGGGCGAGTGCCAT 318
 QY 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60
 DB 319 GCAGACGCGCAATGTGTGACCTCCATCTCCAGGATACACTGTGGGTGTTCATCTA 378
 QY 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80
 DB 379 CGCTCCCACTGGGCCAGTATAAGTGTACCTTTGACAAAGCCAGAGAGCGCTGTGCCAAC 438
 QY 81 GluAlaAlaThrMetAlaThrTrpAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis 100
 DB 439 GAAGCTGGGACCATGGCAACCTTACACAGCTCTCTCTATNNCCAGAGGCCAAGTACCAC 498
 QY 101 LeuCysSerAlaGlyTrpLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120
 DB 499 CTGTGCTCAGCAGGCTGGCTGGAGCCGGGTGCTTACCCACAGCCCTTGCCTCC 558
 QY 121 GlnAsnCysGlySerGlyValValGlyIleValAspTyrGlyProArgProAsnLysSer 140
 DB 559 CAGAACTGTGGCTCTGGTGTGGTGGATAGGACTATGACCTAGACCCACACAGAGT 618
 QY 141 GluMetTrpAspValPheCysTyrArgMetLysAspValAsnCysThr***LysValGly 160
 DB 619 GAAATGTGGAGTGTCTTCTGCTATCGGATGAAAGATGTGAATGCACTGCACCTNCAAGTGGC 678

QY 161 TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuLeuGlnValLeuMetSerPhePro 180
 DB 679 TATGTGGAGATGGCTTCTCATACAGTGGAAACCTGTGTGAGGTCTCTGCTTCCCC 738
 QY 181 SerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaAspGlyArg 200
 DB 739 TCACTCAAAACTTCTGACGGAAGTGTGGCTTATTCACACAGCTCAGCTCGAGGCGGT 798
 QY 201 AlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsn 220
 DB 799 GCAITTTCTAGAACACCTGACTGACCTGTCCATCGCGGCACCTCTTTGNCACAGAAC 858
 QY 221 SerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisLeuAlaAsn 240
 DB 859 AGTGGGCTGGGGAGAAATGAGACCTTCTCTGGGCGGACATCGAGACCCACCTCGCAAT 918
 QY 241 ValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGly 260
 DB 919 GTGAGCATGTTTTCTACAAATGACCTTGTCAATGGCACCCCTGCAACGAGGCTGGGA 978
 QY 261 SerLysLeuLeuIleThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal 280
 DB 979 AGCAAGCTGTCTCATCTGACAGACAGACCCCACTCCACCGGACGAGACCGAGTGTGT 1038
 QY 281 AspGlyArgAspThrLeuGluTrpAspIleCysAlaSerAsnGlyIleThrHisValle 300
 DB 1039 GATGGAAGAGACACTCTGGAGTGGGACATCTGTGCTCCATGGATGCACATGTCTATT 1098
 QY 301 SerArg***LeuLysAlaProProAlaProValThrLeu***HisThrGlyLeuGly*** 320
 DB 1099 TCCAGGYCTTTAAAGCACCCCTGCCCCGCTGACCTTGNCCACACTGNTTGGAGNA 1158
 QY 321 GlyIlePhe*****IleLeuValThrGlyAlaValAlaLeuAlaTyrSerTyr 340
 DB 1159 GGGATCTTCTGNCATCATCTGCTGACTGGGCTGTGCTTGGCTTACTTCTCTTAC 1218
 QY 341 PheArgIleAsnArgLysThrIleGlyPhe***HisPhe 353
 DB 1219 TTTCGGATAAACCGGAAACAAATCGGCTTCCANCAATTT 1257
 RESULT 2
 ABA04662
 ID ABA04662 standard; cDNA; 4576 BP.
 XX AC ABA04662;
 XX DT 22-FEB-2002 (first entry)
 XX DE Human Hyaluronic Acid Receptor for Endocytosis, HARE, coding sequence.
 XX KW HARE; Hyaluronic Acid Receptor for Endocytosis; hyaluronan; chondroitin;
 XX KW chondroitin sulphate; extracellular matrix; cartilage; skin;
 XX OS vitreous humour; endocytic receptor; glycosaminoglycan; human; ss.
 XX Key Location/Qualifiers
 FT CDS 1..4185
 FT /*tag= a
 FT /partial
 FT /product= "Human HARE"
 FT /note= "No start codon given"
 XX WO200181544-A2.
 XX 01-NOV-2001.
 XX 25-APR-2001; 2001WO-US013403.
 XX 25-APR-2000; 2000US-0199538P.
 XX 02-NOV-2000; 2000US-0245320P.
 XX

PA (WEIG/) WEIGEL P A.
 PA (ZHOU/) ZHOU B.
 XX (WEIG/) WEIGEL J A.
 XX Weigel PA, Zhou B, Weigel JA;
 DR WPI; 2002-049271/06.
 DR P-PSDB; AAM47684.
 XX New mammalian hyaluronic acid receptor for endocytosis, useful e.g. for
 PT identifying agents that inhibit binding to hyaluronic acid, and related
 PT nucleic acid.
 XX Claim 49; Fig 33; 263pp; English.
 XX The present invention relates to sequences for rat and human HARE
 CC (Hyaluronic Acid Receptor for Endocytosis, ABA04648, ABA04662, AAM47675
 CC and AAM47684). HARE can bind specifically to at least one of hyaluronic
 CC acid (HA, also known as hyaluronan), chondroitin (C) or chondroitin
 CC sulphate (CS). HA is an extracellular matrix component of all tissues, in
 CC particular cartilage, skin and vitreous humour. HARE is the endocytic
 CC receptor responsible for removing HA and other glycosaminoglycans from
 CC the circulation. The present sequence is the coding sequence for human
 CC HARE
 XX
 SQ Sequence 4576 BP; 1127 A; 1193 C; 1203 G; 1053 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 2,73e-193 Length: 4576
 Score: 1772.00 Matches: 334
 Percent Similarity: 95.18% Conservative: 2
 Best Local Similarity: 94.62% Mismatches: 17
 Query Match: 95.01% Indels: 0
 DB: Gaps: 0
 US-09-466-778b-11 (1-353) x ABA04662 (1-4576)
 QY 1 MetThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeu 20
 Db 2956 ATGACAGGCGCGGAGCAACAGTGTGAGTGTAAAGTCACTATGTCGGAGATGGGCTG 3015
 QY 21 AspCysGluProGluGlnLeuProIleAspArgCysLeuGlnAspGlnCysHis 40
 Db 3016 AACTGTGAGCGGAGAGAGCTGCCATATGACCTGCTTACGAGCAATGGGCGAGTGGCAT 3075
 QY 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60
 Db 3076 GCAGACGCCAAATGTGTGCGACCTCCACTTCCAGGATACCACTGTGGGGTGTTCATCTA 3135
 QY 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80
 Db 3136 CGCTCCCACTGGGCCAGATATAGCTGACCTTTGACAAAGCCAGAGAGGCTGTGCCAAC 3195
 QY 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis 100
 Db 3196 GAGCTGCGACCATGGCAACTACACCACTCTCTATCCCGAGAGGCCAAGTACCA 3255
 QY 101 LeuCysSerAlaGlyTrpLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120
 Db 3256 CTGTGCTCAGAGCTGGCTGGAGACCGGGCGGTGGCTTACCCACAGCCTTCGCGCTCC 3315
 QY 121 GluAsnCysGlySerGlyValValGlyValAlaAspTyrGlyProArgProAsnLysSer 140
 Db 3316 CAGAACTGTGGCTCTGTGTGGTGGATGTGGACTATGACCTAGACCAACAAGAGT 3375
 QY 141 GluMetTrpAspValPheCysTyrArgMetLysAspValAsnCysThr***LysValGly 160
 Db 3376 GAAATGTGGATGTCTTCTCTATTCGATGAAGATGTGAATGTGAATGTGAATGTGAATGTGA 3435
 QY 161 TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuLeuGlnValLeuMetSerPhePro 180
 Db 3436 TATGTGGAGATGGCTTCTCATGTCAGTGGGAACCTGCTGTCAGGTCTCTGATGCTCTTCCC 3495

QY 181 SerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArg 200
 Db 3496 TCACCTACAAACTTCTGACGGAAGTGGCTATTTCAACAGCTCAGCTCGAGCCGT 3555
 QY 201 AlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsn 220
 Db 3556 GCATTTCTAGACACCTGACTGACCTGCTCATCCGGCGGACCTCTTTGTGCCACAGAAC 3615
 QY 221 SerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisLeuAlaAsn 240
 Db 3616 AGTGGGCTGGGAGAAATCAGACCTTGTCTGGCGGGACATCGAGCACCACTCGCAAT 3675
 QY 241 ValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGly 260
 Db 3676 GTCAGCATCTTTTTCACAAATGACCTTGTCAATGGCCACCACCTGCAACGAGGCTGGGA 3735
 QY 261 SerLysLeuLeuIleThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal 280
 Db 3736 AGCAAGCTGCTCATCTACTGCCAGCCAGGACCCACTCCACCCGAGGAGGAGGAGGTTGTT 3795
 QY 281 AspGlyArgAspThrLeuGluTrpAspIleCysAlaSerAsnGlyIleThrHisValIle 300
 Db 3796 GATGGAAGAGCCATTCTGCAGTGGGACATCTTTCCTCCCAATGGGATCATTCATGTCATT 3855
 QY 301 SerArg***LeuLysAlaProAlaProValThrLeu***HisThrGlyLeuGly*** 320
 Db 3856 TCCAGGCTTTTAAAGCACCCCTGCCCCGCTGACCTTGACCCACACTGGCTGGGAGCA 3915
 QY 321 GlyIlePhe*****IleIleLeuValThrGlyValAlaValAlaLeuAlaTyrSerTyr 340
 Db 3916 GGAATCTTTTGGCCATCATCTGCTGACTGGGGCTGTGGCTTGGCTGCTTACTCTACTAC 3975
 QY 341 PheArgIleAsnArgLysThrIleGlyPhe***HisPhe 353
 Db 3976 TTTGGATAAACCGGAGAAACAATCGGCTTCCAGCATTTT 4014
 RESULT 3
 ABX13822
 ID ABX13822 standard; cDNA; 4642 BP.
 XX
 AC ABX13822;
 DT 19-FEB-2003 (first entry)
 XX
 DE cDNA encoding human 190kDa Hyaluronan receptor for endocytosis #1.
 XX
 KW Hyaluronan receptor for endocytosis; HARE; hyaluronan; HA; CD;
 KW chondroitin molecule; chondroitin sulphate; CDS; targeting; metastasis;
 XX tumour; gene therapy; human; gene; ss.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT 1. .4251
 FT /*tag= a
 FT /product= "HARE"
 FT /note= "Hyaluronan receptor for endocytosis"
 FT /partial
 FT /note= "No start codon given"
 XX
 PN WO200286093-A2.
 XX
 PD 31-OCT-2002.
 XX
 PF 25-APR-2002; 2002WO-US013209.
 XX
 PR 25-APR-2001; 2001US-00842930.
 PR 25-APR-2001; 2001US-0286468P.
 XX
 PA (WEIG/) WEIGEL P H.
 PA (WEIG/) WEIGEL J A.
 XX
 PI Weigel PH, Weigel JA;

XX WPI; 2003-093126/08.
 DR P-PSDB; ABG72499.
 XX
 XX Targeting compounds e.g. chemotherapeutic agent to cell of subject
 PT expressing functional active hyaluronan receptor for endocytosis of HARE,
 PT by using hyaluronan, HARE, and/or monoclonal antibody that binds to HARE
 PT epitope.
 XX
 XX Example; Fig 9A; 167pp; English.
 XX
 CC The invention describes a method of targeting a compound to a cell or
 CC tissue of an individual expressing a functionally active hyaluronan (HA)
 CC receptor for endocytosis (HARE) or a cell that does not express
 CC functionally active HARE. The method involves using HA molecule, a
 CC chondroitin molecule (CD), a chondroitin sulphate (CDS) molecule, and/or
 CC a monoclonal antibody raised against a HA-binding domain of HARE. The
 CC method is useful for targeting a compound, preferably a chemotherapeutic
 CC agent or a radioisotope to cell of an individual, especially a human,
 CC expressing HARE on its surface (e.g. gene therapy). Also described is a
 CC method useful for preventing interaction between a cell having at least
 CC one of a HA coat, a CD coat and CDS coat and a cell expressing HARE on
 CC its surface. This second method is useful for preventing metastasis by
 CC preventing interaction between tumour cells having HA, CD or CDS coat and
 CC non-tumour cells expressing HARE on its surface. The invention also
 CC describes a method useful for detecting the presence of HA, CD and CDS in
 CC a biological fluid. This sequence encodes the human 190kDa Hyaluronan
 CC receptor for endocytosis (HARE)
 XX
 SQ Sequence 4642 BP; 1141 A; 1215 C; 1216 G; 1070 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 2,79e-193 Length: 4642
 Score: 1772.00 Matches: 334
 Percent Similarity: 95.18% Conservative: 2
 Best Local Similarity: 94.62% Mismatches: 17
 Query Match: 95.01% Indels: 0
 DB: 7 Gaps: 0
 US-09-466-778b-11 (1-353) x ABX13822 (1-4642)

QY 161 TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuLeuGlnValLeuMetSerPhePro 180
 DB 3502 TATGTGGAGATGGCTTCTCATGAGTGGAACTGCTGAGGTCTCTGATGTCTCTCC 3561
 QY 181 SerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArg 200
 DB 3562 TCACCTCAAACTTCTGACGGAAGTGTGGCTATTCCAAACAGCTCAGCTCAGGCGCGT 3621
 QY 201 AlaPheLeuGluHisLeuThrAspLeuSerLeuArgGlyThrLeuPheValProGlnAsn 220
 DB 3622 GCATTTCTAGAACACCTGACTGACCTGTCCATCGGGCCACCTCTTTGTGCCACAGAAC 3681
 QY 221 SerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspLeuGluHisLeuAlaAsn 240
 DB 3682 AGTGGGCTGGGAGAAATGAGACTTGTCTGGGGGGGACATCGAGCACCACTCGCCAAT 3741
 QY 241 ValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGly 260
 DB 3742 GTCAGCATGTTTTTCTACATGACCTGTCAATGGCACCACTTGCCTGCAACAGAGCTGGGA 3801
 QY 261 SerLysLeuLeuLeuThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal 280
 DB 3802 AGCAAGCTGCTCATCTGTCAGCAGGAGCCCACTCCAAACCGAGAGACCACTGTTGT 3861
 QY 281 AspGlyArgAspThrLeuGluTrpAspPheCysAlaSerAsnGlyIleThrHisValIle 300
 DB 3862 GATGGAAGAGCCCAATTCGACGTGGGACATCTTGGCTTCCAAATGGGATCATTCATGTCAT 3921
 QY 301 SerArg***LeuLysAlaProAlaProValThrLeu***HisThrGlyLeuGly*** 320
 DB 3922 TCCAGGCCCTTTAAAGACACCCCTGCCCGCTGACCTTGACCCACACTGGCTGGAGCA 3981
 QY 321 GlyIlePhe*****IleLeuValThrGlyAlaValAlaLeuAlaLysTyrSerTyr 340
 DB 3982 GGGATCTTCTTCCCATCATCTGCTGAGTGGGGCTGTTCCTTGGCTGCTTACTCTCTAC 4041
 QY 341 PheArgIleAsnArgLysThrIleGlyPhe***HisPhe 353
 DB 4042 TTTCCGATAAACCGAGAGACATCGGCTTCCAGCATTTT 4080
 RESULT 4
 ABX13823
 ID ABX13823 standard; cDNA; 4962 BP.
 XX
 AC ABX13823;
 XX
 DT 19-FEB-2003 (first entry)
 XX
 DE cDNA encoding human 190kDa Hyaluronan receptor for endocytosis #2.
 XX
 KW Hyaluronan receptor for endocytosis; HARE; hyaluronan; HA; CD;
 KW chondroitin molecule; chondroitin sulphate; CDS; targeting; metastasis;
 KW tumour; gene therapy; human; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1. 4962
 FT /*tag= a
 FT /product= "HARE"
 FT /note= "Hyaluronan receptor for endocytosis"
 FT /partial
 FT /note= "No start codon given"
 XX
 PN WO200286093-A2.
 XX
 PD 31-OCT-2002.
 XX
 PF 25-APR-2002; 2002WO-US013209.
 XX
 PR 25-APR-2001; 2001US-00842930.
 PR 25-APR-2001; 2001US-0286468P.
 XX

PA (WEIG/) WEIGEL P H.
 PA (WEIG/) WEIGEL J A.
 PI Weigel PH, Weigel JA;
 XX WPI: 2003-093126/08.
 DR P-PSDB; ABG72514.
 XX
 Targeting compounds e.g. chemotherapeutic agent to cell of subject
 PT expressing functional active hyaluronan receptor for endocytosis of HARE,
 PT by using hyaluronan, HARE, and/or monoclonal antibody that binds to HARE
 PT epitope.
 XX
 PS Example; Fig 9B; 167pp; English.
 XX
 CC The invention describes a method of targeting a compound to a cell or
 CC tissue of an individual expressing a functionally active hyaluronan (HA)
 CC receptor for endocytosis (HARE) or a cell that does not express
 CC functionally active HARE. The method involves using HA molecule, a
 CC chondroitin molecule (CD), a chondroitin sulphate (CDS) molecule, and/or
 CC a monoclonal antibody raised against a HA-binding domain of HARE. The
 CC method is useful for targeting a compound, preferably a chemotherapeutic
 CC agent or a radioisotope to cell of an individual, especially a human,
 CC expressing HARE on its surface (e.g. gene therapy). Also described is a
 CC method useful for preventing interaction between a cell having at least
 CC one of a HA coat, a CD coat and CDS coat and a cell expressing HARE on
 CC its surface. This second method is useful for preventing metastasis by
 CC preventing interaction between tumour cells having HA, CD or CDS coat and
 CC non-tumour cells expressing HARE on its surface. The invention also
 CC describes a method useful for detecting the presence of HA, CD and CDS in
 CC a biological fluid. This sequence encodes a longer version of the human
 CC 190kDa Hyaluronan receptor for endocytosis (HARE) shown in ABG72499
 CC (Encoded by ABX13822)
 XX
 SQ Sequence 4962 BP; 1249 A; 1277 C; 1311 G; 1125 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 3,08e-193 Length: 4962
 Score: 1772.00 Matches: 334
 Percent Similarity: 95.18% Conservative: 2
 Best Local Similarity: 94.62% Mismatches: 17
 Query Match: 95.01% Indels: 0
 DB: 7 Gaps: 0
 US-09-466-778b-11 (1-353) x ABX13823 (1-4962)
 QY 1 MetThrcGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeu 20
 DB 3733 ATGACAGCCCGGGCAAGCACAAGTGTGAGTGTAAAGTCACTATGTCGGAGATGGGCTG 3792
 QY 21 AsnCysGluProGluGlnLeuProIleAspArgCysLeuGlnAspGlnCysHis 40
 DB 3793 AACTGTAGCGCGGAGCGAGCTGCCCATTTGACCGCTTACAGGACATGGGCGAGTGGCAT 3852
 QY 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60
 DB 3853 GCAGACGCCAAATGTGTGACCTCCATTCAGATACCACTGTGGGGTGTTCATCTA 3912
 QY 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80
 DB 3913 CGTCCCACTGGGCCAGTATAGCTGACCTTTGACAAAGCCAGAGAGGCGCTGTGCCAAC 3972
 QY 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis 100
 DB 3973 GAAGCTCGGACCATGGCAACTCAACACAGCTCTCTATGCCAGAGGCCAAGTATCCAC 4032
 QY 101 LeuCysSerAlaGlyTyrLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120
 DB 4033 CTGTGCTCAGACGGCTGGTGGAGACCGGGGGTGTCCCTACCCACAGCTTCGCGCTCC 4092
 QY 121 GlnAsnCysGlySerGlyValValGlyIleValAspTyrGlyProArgProAsnLysSer 140
 DB 4093 CAGAACTGTGGCTCTGGTGTGGTGGATAGTGGACTATGGACTATGGACTATGGACTATGGAGT 4152

QY 141 GluMetTyrAspValPheCysTyrArgMetLysAspValAsnCysThr***LysValGly 160
 DB 4153 GAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACCTGCACCTGCAGGTGGC 4212
 QY 161 TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuLeuGlnValLeuMetSerPhePro 180
 DB 4213 TATGTGGAGATGGCTTCTCATGCACTGGGAACCTGCTGAGGTCTCTGATGCTCTCC 4272
 QY 181 SerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArg 200
 DB 4273 TCACCTACAACCTCTGACGGAGTGTGCTCTATTCACACAGCTCAGCTCAGAGGCGGT 4332
 QY 201 AlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsn 220
 DB 4333 GCATTTCTAGACACCTGACTGACCTGTCATCCGGGGCACCTCTTTGTGCCACAGAAC 4392
 QY 221 SerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisLeuAlaAsn 240
 DB 4393 AGTGGGCTGGGGAGATGAGACCTTGTCTGGGGGGGACATCGAGCACCACTCGCCAAAT 4452
 QY 241 ValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGly 260
 DB 4453 GTCAGCATGTTTCTACATGACCTGTCAATGGCACCACTCCAAACGAGAGCAGGTGTGTT 4512
 QY 261 SerLysLeuLeuLeuThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal 280
 DB 4513 AGCAAGCTGCTCATCTGTCAGCGAGGAGCCACTCCAAACGAGAGCAGGTGTGTT 4572
 QY 281 AspGlyArgAspThrLeuGluTyrAspIleCysAlaSerAsnGlyIleThrHisValIle 300
 DB 4573 GATGGAGAGCCATCTGTCAGTGGGACATCTTGGCTCCCAATGGGATCATTCATGTCATT 4632
 QY 301 SerArg***LeuLysAlaProProAlaProValThrLeu***HisThrGlyLeuGly*** 320
 DB 4633 TCCAGGCTTTAAAGCAGCCCTGCGCCCGTGCACCTTGACCCACACCTGGCTGGAGCA 4692
 QY 321 GlyIlePhe*****IleLeuValThrGlyAlaValAlaLeuAlaTyrSerTyr 340
 DB 4693 GGGATCTCTTTGCCCATCATCTGCTGCTGCTGGGCTGTGGCTGCTGCTTACTCTCTAC 4752
 QY 341 PheArgIleAsnArgLysThrIleGlyPhe***HisPhe 353
 DB 4753 TTTCCGATAAACCGAGAACATCGGCTTCCAGCATTTT 4791
 RESULT 5
 ABT08489
 ID ABT08489 standard; cDNA; 3260 BP.
 XX
 AC ABT08489;
 XX
 DT 28-NOV-2002 (first entry)
 XX
 DE Human novel protein NOV1c coding sequence SEQ ID NO: 210.
 XX
 KW Human; NOVX; single nucleotide polymorphism; SNP; anti-HIV; cytostatic;
 KW antiarteriosclerotic; antidiabetic; antiasthmatic; antiinflammatory;
 KW haemostatic; hypotensive; neuroprotective; anorectic; nootropic;
 KW antidepressant; immunosuppressive; antibacterial; antiparasitic;
 KW virucide; tranquilizer; anticonvulsant; osteopathic; analgesic;
 KW antiparkinsonian; dermatological; antifertility; cerebroprotective;
 KW antiaddictive; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PH Key Location/Qualifiers
 FT variation /tag= a
 FT variation /tag= b
 FT variation /tag= c

PN W0200259315-A2.
 XX PD 01-AUG-2002.
 XX PF 19-DEC-2001; 2001WO-US0500076.
 XX PR 19-DEC-2000; 2000US-0256619P.
 XX PR 19-JAN-2001; 2001US-0262959P.
 XX PR 28-FEB-2001; 2001US-0272408P.
 XX PR 20-APR-2001; 2001US-0285189P.
 XX PR 26-JUL-2001; 2001US-0308039P.
 XX PR 09-AUG-2001; 2001US-0311266P.
 XX (CURA-) CUPAGEN CORP.
 PA Shinkets RA, Patturajan M, Vernet CAM, Casman SU, Malyankar U;
 XX Shenoy S, Spytek KA, Gangolli E, Miller C, Boldog F, Li L;
 PI Taupier RJ, Kekuda R, Smithson G, Zerhusen BD, Liu X, Colman SD;
 PI Tchernev V, Si J, Edinger S, Stone D, Sciore P, Millet I;
 PI Rothenberg M;
 XX WPI; 2002-666903/71.
 DR P-PSDB; ABJ10588.
 XX
 PT New isolated NOVX polypeptides and polynucleotides, useful for
 PT preventing, diagnosing or treating NOVX-associated disorders e.g.
 PT diabetes, Crohn's disease, atherosclerosis, cancer, Huntington's disease
 PT or Alzheimer's disease.
 XX
 PS Claim 42; Page 24-25; 363pp; English.
 CC
 CC The present invention provides the protein and coding sequences of
 CC several novel human proteins, designated NOVX. These can be used in the
 CC treatment of diseases such as cancers, Hodgkin's disease, Von Hippel-
 CC Lindau syndrome, Alzheimer's disease, stroke, tuberosus sclerosis,
 CC hypercalcaemia, Parkinson's disease, Huntington's disease, cerebral
 CC palsy, epilepsy, Lesch-Nyhan syndrome, multiple sclerosis, ataxia
 CC telangiectasia, leukodystrophies, addiction, anxiety, depression, pain,
 CC obesity, Crohn's disease, osteoporosis, inflammatory bowel disease,
 CC infertility, atherosclerosis, hypertension, scleroderma, haemophilia,
 CC diabetes, pancreatitis, autoimmune disease, asthma, arthritis,
 CC immunodeficiencies, HIV, viral, bacterial or parasitic infections, or
 CC graft-versus-host disease. The present sequence is a coding sequence of
 CC the invention
 XX
 SQ Sequence 3260 BP; 774 A; 886 C; 850 G; 749 T; 0 U; 1 Other;
 Alignment Scores:
 Pred. No.: 3,678-193 Length: 3260
 Score: 1769.00 Matches: 333
 Percent Similarity: 95.18% Conservative: 3
 Best Local Similarity: 94.33% Mismatches: 17
 Query Match: 94.85% Indels: 0
 DB: 6 Gaps: 0
 US-09-466-778B-11 (1-353) x ABT08489 (1-3260)
 QY 1 MetThrGlyProGlyValHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeu 20
 DB 1635 ATGACAGCGCCGGGCAAGCAAGGTGAGTGAAGTCACTATGTGCGAGATGGGCTG 1694
 QY 21 AsnCysGluProGluGlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis 40
 DB 1695 AACTGTGAGCGCGAGACAGCTGCCATTCACCTGCTTACAGGACAAATGGGCGATGCCAT 1754
 QY 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60
 DB 1755 GCAGACGCCAAATGTGTGACCTCCACCTCCAGGATACCACTGTGGGGTGTTCATCTA 1814
 QY 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80
 DB 1815 CGCTCCCACTGGGCGCAGTATAAGTACCTTTGCAAGGCCAGAGAGCGCTGTGCCAAC 1874

QY 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis 100
 DB 1875 GAAGCTGGACCATGGCAACCTACCAACAGCTCTCTATGCCAGAGCCCAAGTACCAAC 1934
 QY 101 LeuCysSerAlaGlyTyrLeuGluThrGlyArgValAlaLysProThrAlaPheAlaSer 120
 DB 1935 CTGTGCTCAGCAGGCTGGCTGGAGCGGGGGGTGGCTTACCCACACAGCTTCGCCTCC 1994
 QY 121 GlnAsnCysGlySerGlyValValGlyLysValAspTyrGlyProArgProAsnLysSer 140
 DB 1995 CAGAACTGTGGCTCTGGTGTGGTGGGATAGTGGACTATGACCTAGACCCCAACAGAGT 2054
 QY 141 GluMetTyrAspValPheCysTyrArgMetLysAspValAsnCysThr***LysValGly 160
 DB 2055 GAAATGTGGGATGTCTTCTGTCTATCGATGAAGATGTGAACCTGACCTGCAAGGTGGC 2114
 QY 161 TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuGlnValLeuMetSerPhePro 180
 DB 2115 TATGTGGAGATGGCTTCTCATGAGTGGGAACTGTGCGAGTCTCTGATCTCTTCCCC 2174
 QY 181 SerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArg 200
 DB 2175 TCACCTCAAACTTCTCTGACGGAAGTGTCTGCTTATTCACACAGCTCAGCTCGAGCCGT 2234
 QY 201 AlaPheLeuGluHisLeuThrAspLeuSerTyrLeuArgGlyThrLeuPheValProGlnAsn 220
 DB 2235 GCATTTCTAGAACACCTGACCTGACCTGCTCCATCCGCGGCACTCTTTGTGCCACAGAAC 2294
 QY 221 SerGlyLeuGlyGluAsnGluThrLeuSerGlyValArgAspLeuHisLeuAlaAsn 240
 DB 2295 AGTGGCTGGGGAGATGAGACCTTGTCTGGCGGGAGACATCGAGACCCACCTCGCCAT 2354
 QY 241 ValSerMetPheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGly 260
 DB 2355 GTACGATGTTTTTCTCAATGACCTTGTCAATGGCACCACCTGCAAAACGAGGGTGGGA 2414
 QY 261 SerLysLeuLeuLeuThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal 280
 DB 2415 AGCAAGTGTCTCATCTGCGCGAGCCAGCCACCTCCAAACGAGGAGACCAAGTTGTT 2474
 QY 281 AspGlyArgAspThrLeuGluTyrAspLeuCysAlaSerAsnGlyIleThrHisValIle 300
 DB 2475 GATGGAGAGCCATCTCTGACGTGGGACATCTTTCCCTCCAAATGGGATCATTCATGCT 2534
 QY 301 SerArg***LeuLysAlaProProAlaProValThrLeu***HisThrGlyLeuGly*** 320
 DB 2535 TCCAGGCTTTTAAAGACACCCCTGCGCCCTGACCTTGACCCACACTGGCTGGGAGCA 2594
 QY 321 GlyIlePhe*****IleLeuValThrGlyAlaValAlaLeuAlaLysTyrSerTyr 340
 DB 2595 GGGATCTTTTGGCATCATCTGCTGACTGGGGCTGTGCTGCTGGCTGCTTACTCTCTAC 2654
 QY 341 PheArgIleAsnArgLysThrIleGlyPhe***HisPhe 353
 DB 2655 TTTCCGATAAACCGGAGACATCGGCTTCCAGCATTTT 2693
 RESULT 6
 AAF87120
 ID AAF87120 standard; DNA; 3625 BP.
 XX
 AC AAF87120;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE NOV9 coding sequence.
 XX
 KW NOV; Cytostatic; contraceptive; antiinflammatory; immunomodulatory;
 KW cardiovascular; casein kinase II phosphorylation site; contraction;
 KW serine/threonine kinase; Peutz-Jeghers syndrome; cellular proliferation;
 KW epidermal growth factor; cell development; apoptosis; cell adhesion;
 KW growth migration; cell structure; motility; cancer; immune disorder;
 KW inflammatory disorder; cellular adhesion disorder; long-QT syndrome;
 KW cardiovascular disease; hypertrophic cardiomyopathy; marfan syndrome;

KW therapy; NOV1; NOV2; NOV3; NOV4; NOV5; NOV6; NOV7; NOV8; NOV9; NOV10;
 KW NOV11; NOV12; NOV13; NOV14; NOV15; NOV16; ds.
 OS Unidentified.

XX Key Location/Qualifiers
 XX CDS 387..2024
 FT /*tag= a
 FT /product= "NOV9"

XX WO200136638-A2.

XX 25-MAY-2001.

XX 17-NOV-2000; 2000WO-US0311543.

XX 19-NOV-1999; 99US-0166336P.

XX 29-NOV-1999; 99US-0167785P.

XX 08-MAR-2000; 2000US-0187844P.

XX 16-NOV-2000; 2000US-00715417.

XX (CURA-) CURAGEN CORP.

XX Shimkets RA, Lichenstein H, Vernet C, Fernandes E;

XX WFI; 2001-648134/74.

XX P-PSDB; AAB83365.

XX Novel human polypeptides and the nucleic acids that encode them useful
 PT for preventing, diagnosing and treating e.g. cancer, inflammation and
 PT immune disorders.

XX Claim 9; Page 32-33; 141pp; English.

XX This sequence encodes the NOV9 protein. The invention relates to the NOV1
 CC -NOV16 proteins, and their coding sequences. The proteins have Cytostatic
 CC ; contraceptive; antiinflammatory; immunomodulatory; and cardiovascular
 CC activities. The sequences may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate NOVX expression. They
 CC may be used to treat disorders associated with decreased expression by
 CC rectifying mutations or deletions in a patient's genome that affect the
 CC activity of protein by expressing inactive proteins or to supplement the
 CC patients own production of protein. They are used to produce NOVX
 CC proteins, by inserting the nucleic acid into a cell and culturing it to
 CC express the protein. The DNA may be used as DNA probes in assays to
 CC detect and quantitate the presence of similar DNAs in samples, and which
 CC patients may need restorative therapy. The NOVX protein may also be used
 CC as antigens in the production of antibodies (Abs) against NOVX and in
 CC assays to identify modulators of NOVX expression and activity. The anti-
 CC NOVX Abs and antagonist are used to down regulate expression and
 CC activity. The anti-NOVX Abs are used for detecting the presence of NOVX
 CC in samples. Disorders that may be prevented, diagnosed and/or treated
 CC vary depending on the NOVX protein. NOV1, NOV3, NOV5, NOV7, NOV9-11 and
 CC NOV13-16 have casein kinase II phosphorylation sites characteristic of
 CC serine/threonine kinases, and are used to treat kinase-related disorders
 CC (e.g. Peutz-Jeghers syndrome, cellular proliferation and contraception).
 CC NOV2-5, NOV6 and NOV8 are homologous to the epidermal growth factor (EGF).
 CC -like super family and are involved in, e.g. regulation of cell
 CC development, apoptosis, cell adhesion, growth migration, cell structure
 CC and motility and protein management, and are used to treat cancers,
 CC inflammatory disorders, immune disorders and cellular adhesion disorders.
 CC NOV6-10 are homologous to EGF-like fibrillin proteins and are used to
 CC treat cardiovascular disease e.g. hypertrophic cardiomyopathy, long-Qr
 CC syndrome and marfan syndrome

XX SQ Sequence 3625 BP; 896 A; 971 C; 914 G; 844 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 9.58e-169 Length: 3625
 Score: 1558.00 Matches: 300
 Percent Similarity: 85.31% Conservative: 2
 Best Local Similarity: 84.75% Mismatches: 15
 Query Match: 83.54% Indels: 37

DB:	4	Gaps:	1
US-09-466-778B-11 (1-353) x AAF87120 (1-3625)			
Qy	1	MetThrGlyProGlyLysHisLysCysGluCysAlaSerHisTyrValGlyAspGlyLeu	20
Db	2078	ATGACAGGCGCGGCAAGCAAGTGTGAGTGTAAAGTCACTATGTCGAGATGGCTG	2137
Qy	21	AsnCysGluProGluGlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis	40
Db	2138	AACTGTGAGCGGAGCAGCTGCCAATTGACCGTCTTACAGACATGGCAGTGCAT	2197
Qy	41	AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu	60
Db	2198	GCAGAGCCCAATGTGTGACCTCCACTTCAGGATACCACTGTGGGTGTTCATCTA	2257
Qy	61	ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn	80
Db	2258	CGCTCCCACTGGGCCAGTATAGCTGACCTTTGACAAAGCCAGAGAGGCTGTGCCAAC	2317
Qy	81	GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis	100
Db	2318	GAAGCTGGACCATGGCAACCTACACCACTCTCTATGCCAGAGGCCAGTACCAC	2377
Qy	101	LeuCysSerAlaGlyTrpLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer	120
Db	2378	CTGTGCTCAGAGGCTGGCTGGAGACGGCGGGGTGGCTTACCCACACAGCCTTCGCCCTCC	2437
Qy	121	GlnAsnCysGlySerGlyValValGlyIleValAspTyrGlyProArgProAsnLysSer	140
Db	2438	CAGAACTGTGGCTCTGTGTGGTGGATAGTGGACTATGGACTAGACCACCAACAGAGT	2497
Qy	141	GluMetTrpAspValPheCysTyrArgMet - LysAspValAsnCysThr***LysValG1	160
Db	2498	GAATGTGGGATGTCTTCGCTATCGGATGAAAG-----	2531
Qy	160	YTrValGlyAspGlyPheSerTyrSerGlyAsnLeuLeuGlnValLeuMetSerPhePr	180
Db	2531	-----	2531
Qy	180	oSerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyAr	200
Db	2532	-----GAAGTGTGGGCTATTTCCACAGCTCACTCGAGGCGG	2569
Qy	200	GAlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAs	220
Db	2570	TGCATTCTAGAACACCTGACTGACCTGCTCCATCCGCGGCACCTCTTTTGTGCCACAGAA	2629
Qy	220	nSerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisLeuAlaAs	240
Db	2630	CAGTGGGCTGGGGAGAAATGAGACCTTGTCTGGCGGGAGACATCGAGCACCACCTCGCAA	2689
Qy	240	nValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuG1	260
Db	2690	TGTCAGCATGTTTTCTACATGACCTTGTCAATGGCACCCCTGCAACAGAGGCTGGG	2749
Qy	260	YSerLysLeuLeuIleThrAspArgGlnAspProLeuHisProThrGluThrArgCysVa	280
Db	2750	AAGCAAGCTGCTCATCTACCTGCCAGCCAGGACCCCACTCCACCGCAGGAGACAGTTGT	2809
Qy	280	lAspGlyArgAspThrLeuGluTrpAspIleCysAlaSerAsnGlyIleThrHisVal11	300
Db	2810	TGATGAAGAGGCCATCTTCGAGTGGGACATCTTTGCTCCAAATGGGATCATTCATGTCAT	2869
Qy	300	eSerArg***LeuLysAlaProProAlaProValThrLeu***HisThrGlyLeuGly**	320
Db	2870	TTCCAGGCTTTAAAGACCCCTCCCGCGGACCTTGACCCACACCTGGCTGGGAGC	2929
Qy	320	*GlyIlePhe*****IleIleLeuValThrGlyAlaValAlaLeuAlaAlaTyrSerTy	340
Db	2930	AGGGA1CTTTTGGCCATCATCTGCTGGTGGGCTGTTGCTTGGCTTGGCTTACTCTTA	2989
Qy	340	rPheArgIleAsnArgLysThrIleGlyPhe***HisPhe	353

Db 2990 CTTTCGGATAACCGGAGCAATCGGCTCCAGCATTTT 3029

RESULT 7

ID ABT08488 standard; cDNA; 8495 BP.

AC ABT08488;

DT 28-NOV-2002 (first entry)

XX Human novel protein NOV1b coding sequence SEQ ID NO: 3.

XX Human; NOVX; single nucleotide polymorphism; SNP; anti-HIV; cytostatic;
 KW antiarteriosclerotic; antidiabetic; antiasthmatic; antiinflammatory;
 KW haemostatic; hypotensive; neuroprotective; anorectic; nootropic;
 KW antidepressant; immunosuppressive; antibacterial; antiparasitic;
 KW virucide; tranquilizer; anticonvulsant; osteopathic; analgesic;
 KW antiparkinsonian; dermatological; antiinfertility; cerebroprotective;
 KW antiaddictive; gene; ss.

XX Homo sapiens.

OS WO200259315-A2.

PN 01-AUG-2002.

XX 19-DEC-2001; 2001WO-US050076.

XX 19-DEC-2000; 2000US-0256619P.

XX 19-JAN-2001; 2001US-0262959P.

XX 28-FEB-2001; 2001US-0272408P.

XX 20-APR-2001; 2001US-0285189P.

XX 26-JUL-2001; 2001US-0308039P.

XX 09-AUG-2001; 2001US-0311266P.

XX (CURA-) CURAGEN CORP.

XX Shimkets RA, Patturajan M, Vernet CAM, Casman SJ, Malyankar U;

XX Shenoy S, Spytek KA, Gangolli E, Miller C, Boldog F, Li L;

XX Taupier RJ, Kekuda R, Smithson G, Zerhusen BD, Liu X, Colman SD;

XX Tchernev V, Si J, Edinger S, Stone D, Sciore P, Millet I;

XX Rothenberg M;

XX WPI; 2002-666903/71.

XX P-FSDB; ABJ10587.

XX New isolated NOVX polypeptides and polynucleotides, useful for
 PT preventing, diagnosing or treating NOVX-associated disorders e.g.
 PT diabetes, Crohn's disease, atherosclerosis, cancer, Huntington's disease
 or Alzheimer's disease.

XX Claim 8; Page 21-23; 363pp; English.

XX The present invention provides the protein and coding sequences of
 CC several novel human proteins, designated NOVX. These can be used in the
 CC treatment of diseases such as cancers, Hodgkin's disease, Von Hippel-
 CC Lindau syndrome, Alzheimer's disease, stroke, tuberous sclerosis,
 CC hypercalcaemia, Parkinson's disease, Huntington's disease, cerebral
 CC palsy, epilepsy, Lesch-Nyhan syndrome, multiple sclerosis, ataxia
 CC telangiectasia, leukodystrophies, addiction, anxiety, depression, pain,
 CC obesity, Crohn's disease, osteoporosis, inflammatory bowel disease,
 CC infertility, atherosclerosis, hypertension, scleroderma, haemophilia,
 CC diabetes, pancreatitis, autoimmune disease, asthma, arthritis,
 CC immunodeficiencies, HIV, viral, bacterial or parasitic infections, or
 CC graft-versus-host disease. The present sequence is a coding sequence of
 the invention

XX SQ Sequence 8495 BP; 2295 A; 2102 C; 2135 G; 1957 T; 0 U; 6 Other;

Alignment Scores:

Pred. No.: 3.39e-168 Length: 8495

Score: 1558.00 Matches: 300

Percent Similarity: 85.31% Conservative: 2
 Best Local Similarity: 84.75% Mismatches: 15
 Query Match: 83.54% Indels: 37
 DB: 6 Gaps: 1

US-09-466-778B-11 (1-353) x ABT08488 (1-8495)

QY	1	MetThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeu	20
Db	6948	ATGACAGGCCCGGCAAGACCAAGTGTGAGTAAAGTCACTATGTCGAGATGGCTG	7007
QY	21	AsnCysGluProGluGlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis	40
Db	7008	AACTGTAGCCGAGCAGCAGTGCCTTACAGGACATGGCGAGTGGCAT	7067
QY	41	AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu	60
Db	7068	GCAGAGCCCAATGTGTCCAGCTCCACTTCAGATACCACTGTGGGGTGTTCATCTA	7127
QY	61	ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn	80
Db	7128	CGCTCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCTGTGCCAAC	7187
QY	81	GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr**GlnLysAlaLysTyrHis	100
Db	7188	GAAGCTCGACCATGGCACTCAACACCTCTCTATGCCAGAAAGGCCAAGTACCA	7247
QY	101	LeuCysSerAlaGlyTrpLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer	120
Db	7248	CTGTGCTCAGCAGCTGGCTGGAGACCGGGCGGGTTGCCTACCCACAGCCTTCGCGCTCC	7307
QY	121	GlnAsnCysGlySerGlyValValGlyIleValAspTyrGlyProArgProAsnLysSer	140
Db	7308	CAGAACTGTGGCTCTGGTGTGGATAGTGGACTATGGACCTAGACCAACAGAGT	7367
QY	141	GluMetTrpAspValPheCysTyrArgMet-LysAspValAsnCysThr**LysValG1	160
Db	7368	GAATGTGGATGTCTTCTGCTATCGATGAAG-----	7401
QY	160	yTyrValGlyAspGlyPheSerTyrSerGlyAsnLeuLeuGlnValLeuMetSerPhePr	180
Db	7401	-----	7401
QY	180	oSerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyAr	200
Db	7402	-----GAAAGTGTGGCTATTCCAAAGCTCAGCTCAGGCGCG	7439
QY	200	gAlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAs	220
Db	7440	TGATTTCTAGAACACCTGACTGCTCCATCCCGGCGACCCCTCTTTGTGCCACAGAA	7499
QY	220	nSerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisLeuAlaAs	240
Db	7500	CAGTGGGCTGGGGAGAATGAGACCTTGTCTGGCGGGACATCGAGCACCCCTGCCAA	7559
QY	240	nValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuG1	260
Db	7560	TGTCAGCATGTTTTCTTACAATGACCTTGTCATGGCACCCCTGCAACAGAGGTGGG	7619
QY	260	ySerLysLeuLeuIleThrAspArgGlnAspProLeuHisProThrGluThrArgCysVa	280
Db	7620	AAGCAAGCTGCTCATCTACGCCAGCAGGACCCCTCCAAACCGAGGACCAAGTTGT	7679
QY	280	lAspGlyArgAspThrLeuGluTrpAspIleCysAlaSerAsnGlyIleThrHisVal11	300
Db	7680	TGATGGAAGAGCCATTCTCGAGTGGGACATCTTTGCTCCATGGGATCATTCATGTCA	7739
QY	300	eSerArg***LeuLysAlaProAlaProValThrLeu***HisThrGlyLeuGly**	320
Db	7740	TTCCAGGCTTTAAAGCACCCCTCCCGCGGACCTTGACCCACACTGGCTGGGAGC	7799
QY	320	*GlyIlePhe*****IleIleLeuValThrGlyValAlaValAlaLeuAlaTyrSerTy	340

Db 3859 GATGGAAGATCCATCTGCGAGTGGGACATCATCCCGCCCAATGGAATCCCTCCATATTATT 3918
 QY 301 SerArg***LeuLysAlaProAlaProValThrLeu***HisThrGlyLeuGly*** 320
 Db 3919 TCTGAACCTTTGAGAGTCTCTCCACGGCAGCAAGCGCTGCCACTCTGGCTGGGACACA 3978
 QY 321 GlyIlePhe*****IleLeuValThrGlyAlaValAlaLeuAlaTyrSerTyr 340
 Db 3979 GGTATATTCTGTCGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4038
 QY 341 PheArgIleAsnArgLysThrIleGlyPhe***HisPhe 353
 Db 4039 TTCGGCTAAAGCAGCAGCAACCATCTGGTTTCAGCGGTTT 4077
 RESULT 10
 ABX13821
 ID ABX13821 standard; cDNA; 4706 BP.
 XX
 AC ABX13821;
 XX
 DT 19-FEB-2003 (first entry)
 XX
 DE cDNA encoding rat 175kDa Hyaluronan receptor for endocytosis (HARE).
 XX
 KW Hyaluronan receptor for endocytosis; HARE; hyaluronan; HA; CD;
 KW chondroitin molecule; chondroitin sulphate; CDS; targeting; metastasis;
 KW tumour; gene therapy; rat; gene; ss.
 XX
 OS Rattus sp.
 XX
 FH Key Location/Qualifiers
 FT CDS 1. 4296
 FT /*tag= a
 FT /product= "HARE"
 FT /note= "Hyaluronan receptor for endocytosis"
 FT /partial
 FT /note= "No start codon given"
 XX
 PN WO200286093-A2.
 XX
 PD 31-OCT-2002.
 XX
 PF 25-APR-2002; 2002WO-US013209.
 XX
 PR 25-APR-2001; 2001US-00842930.
 PR 25-APR-2001; 2001US-0286468P.
 XX
 PA (WEIG/) WEIGEL P H.
 PA (WEIG/) WEIGEL J A.
 XX
 PI Weigel PH, Weigel JA;
 XX
 WPI; 2003-093126/08.
 DR P-PSDB; ABG72498.
 XX
 Targeting compounds e.g. chemotherapeutic agent to cell of subject
 PT expressing functional active hyaluronan receptor for endocytosis of HARE,
 PT by using hyaluronan, HARE, and/or monoclonal antibody that binds to HARE
 PT epitope.
 XX
 Example; Fig 2; 167pp; English.
 PS
 CC The invention describes a method of targeting a compound to a cell or
 CC tissue of an individual expressing a functionally active hyaluronan (HA)
 CC receptor for endocytosis (HARE) or a cell that does not express
 CC functionally active HARE. The method involves using HA molecule, a
 CC chondroitin molecule (CD), a chondroitin sulphate (CDS) molecule, and/or
 CC a monoclonal antibody raised against a HA-binding domain of HARE. The
 CC method is useful for targeting a compound, preferably a chemotherapeutic
 CC agent or a radioisotope to cell of an individual, especially a human,
 CC expressing HARE on its surface (e.g. gene therapy). Also described is a
 CC method useful for preventing interaction between a cell having at least

CC one of a HA coat, a CD coat and CDS coat and a cell expressing HARE on
 CC its surface. This second method is useful for preventing metastasis by
 CC preventing interaction between tumour cells having HA, CD or CDS coat and
 CC non-tumour cells expressing HARE on its surface. The invention also
 CC describes a method useful for detecting the presence of HA, CD and CDS in
 CC a biological fluid. This sequence encodes the rat 175kDa Hyaluronan
 CC receptor for endocytosis (HARE)
 XX
 SQ Sequence 4706 BP; 1168 A; 1273 C; 1230 G; 1035 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1,48e-163 Length: 4706
 Score: 1514.50 Matches: 280
 Percent Similarity: 87.54% Conservative: 29
 Best Local Similarity: 79.32% Mismatches: 43
 Query Match: 81.21% Indels: 1
 DB: 7 Gaps: 1
 US-09-466-778B-11 (1-353) x ABX13821 (1-4706)
 QY 1 MetThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeu 20
 Db 3022 ATGACGGGCCAGCGCAAGCATAGTGTGAATGTAAAGTCACTATGTGCGGGACGGAGTG 3081
 QY 21 AsnCysGluProGluGlnLeuProIleAspArgCysLeuGlnAspGlnGlnCysHis 40
 Db 3082 GACTGTGAGCTGAGCAGCTGCCGCTCGACCGTTGTCTTACAGGACACACGACAGTGCAC 3141
 QY 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60
 Db 3142 CCAGATGCCAGCTGTGCAGACCTCTACTCCAGACACAGCCGTAGGAGTATTCATCTA 3201
 QY 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80
 Db 3202 CGTCCCTCCACTGGCCAGTACAACTGACATTTGACAAAGCCAAAGAGCCTGTGCCAAA 3261
 QY 81 GluAlaAlaThrMetalThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis 100
 Db 3262 GAAGCTCGGACCATAGCCACTCAACACAGCTCTCTATGCCAGAGGCAAGTATTCAC 3321
 QY 101 LeuCysSerAlaGlyTrpLeuGluThrGlyValGlyValAlaTyrProThrAlaPheAlaSer 120
 Db 3322 CTGTCTCGCCCGCTGGCTGGAGAGTGGCGGGTGGCTTACCCTACCTACGTATGCTCT 3381
 QY 121 GlnAsnCysGlySerGlyValValGlyIleValAspTyrGlyProArgProAsnLysSer 140
 Db 3382 CAGAAGTGTGGTCAAAACGTTTGTGGATCGTAGACTACGGATCCAGGGCCAAAGAGT 3441
 QY 141 GluMetTrpAspValPheCysTyrArgMetLysAspValAsnCysThr***LysValGly 160
 Db 3442 GAAATGTGGGATGTCTTCTGTACCAGATGAAAGATGTGAACCTGCACCTGCAAGGCGGC 3501
 QY 161 TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuGlnValLeuMetSerPhePro 180
 Db 3502 TATGTGGAGATGGCTTCTGTCAGTGGGAACCTGCTGGAGTCTCTATGCTCTTCCC 3561
 QY 181 SerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArg 200
 Db 3562 TCGCTCACAACTCTCCTGACAGAGGTGTGGCTGTCTTTTCCAAAGAGCTCAGCCCGAGACAG 3621
 QY 201 AlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsn 220
 Db 3622 GCATTTTGAACACCTGACTGACCTGTCCATCCGTGGCCACCTGTTGTGGCCACAGAAC 3681
 QY 221 SerGlyLeuGlyGluAsnGlnLeuThrLeuSerGlyArgAspIleGluHisLeuAlaAsn 240
 Db 3682 AGTGGGCTACCGGGAATTAAGACCTGTCTGGCCGGGACATGAGGACACACCTACTAAT 3741
 QY 241 ValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGly 260
 Db 3742 GTCAACGTCTCTCTTTTACAATGACCTGTCAATGGTACCTTTCTGAGGACTATGCTGGGA 3801
 QY 261 SerLysLeuLeuIleThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal 280

DB 3802 AGCCAACTGCTCATTTACCTTGACGAGGACACAGCTCCAC---CAAGAGACAGGTTTGTG 3858
 QY 281 AspGlyArgAspThrLeuGluTrpAspIleCysAlaSerAsnGlyIleThrHisValIle 300
 DB 3859 GATGGAAGATCCATCTTCGAGTGGGACATCATCGCGGCAATGGAAATCCTCCATATTATT 3918
 QY 301 SerArg***LeuLysAlaProProAlaProValThrLeu***HisThrGlyLeuGly*** 320
 DB 3919 TCTGAACCTTTGAGAGCTCTCCACCGGACGAGCAACGGCTGCCACTCTGGCTGGGGACA 3978
 QY 321 GlyIlePhe*****IleIleLeuValThrGlyAlaValAlaLeuAlaAlaIleTrpSerTyr 340
 DB 3979 GGTATTCTGTGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4038
 QY 341 PheArgIleAsnArgLysThrIleGlyPhe***HisPhe 353
 DB 4039 TTCGGCTAAAGCAGCGAACCACTGGTTCCAGCGTTTT 4077
 RESULT 11
 AAC76373
 ID AAC76373 standard; cDNA; 1377 BP.
 AC AAC76373;
 XX
 DT 08-FEB-2001 (first entry)
 DE Human ORFX ORF1928 polynucleotide sequence SEQ ID NO:3855.
 XX
 KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnery; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antineoplastic;
 KW antitumor; antibacterial; antifungal; antirheumatic; antithyroid;
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200058473-A2.
 XX
 PD 05-OCT-2000.
 XX
 PF 31-MAR-2000; 2000WO-US008621.
 XX
 PR 31-MAR-1999; 99US-0127607P.
 PR 02-APR-1999; 99US-0127636P.
 PR 05-APR-1999; 99US-0127728P.
 PR 30-MAR-2000; 2000US-00540763.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Shinkets RA, Leach M;
 XX
 DR WPI: 2000-602362/57.
 DR P-PDB; AAB42164.
 XX
 PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease.
 XX
 PS Claim 5; Page 3006-3007; 5507pp; English.
 XX
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnery;

CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;
 CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;
 CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
 CC dermatological; immunosuppressive; antineoplastic; antibacterial;
 CC antitumor; antifungal; antirheumatic; antithyroid; antianemic. The
 CC sequences can be used for determining the presence of or predisposition
 CC to or preventing or treating pathological conditions associated with an
 CC ORFX-associated disorder. The nucleic acids can be used to express ORFX
 CC proteins in gene therapy vectors. The proteins and nucleic acids may be
 CC used to treat cancers, proliferative disorders, neurodegenerative
 CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
 CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
 CC storage, systemic lupus erythematosus, severe combined immunodeficiency
 CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
 CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
 CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
 CC enhance coagulation; to inhibit thrombosis; and as a contraceptive
 XX
 SQ Sequence 1377 BP; 303 A; 398 C; 342 G; 330 T; 0 U; 4 Other;

Alignment Scores:

Pred. No.: 1.04e-136 Length: 1377
 Score: 1275.00 Matches: 252
 Percent Similarity: 92.73% Conservative: 3
 Best Local Similarity: 91.64% Mismatches: 17
 Query Match: 68.36% Indels: 3
 DB: 3 Gaps: 0

US-09-466-778B-11 (1-353) x AAC76373 (1-1377)

QY 82 AlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHisLeu 101
 DB 4 GCTGGACCATGGCAACCTCAACACGCTCTCTATGCCAGAGAGGCCAGTACACCTG 63
 QY 102 CysSerAlaGlyTrpLeuGluThrGlyArgValAlaIleTyrProThrAlaPheAlaSerGln 121
 DB 64 TGCTCAGCAGGCTGGCTGGAGACCGGGGGTGGCTACCCACACAGCCTTCGCTCCACG 123
 QY 122 AsnCysGlySerGlyValValGlyIleValAspTyrGlyProArgProAsnLysSerGlu 141
 DB 124 AACTGTGGCTCTGGTGTGGTGTGGATAGTGACATATGGACCTAGACCCCAACAGAGTAA 183
 QY 142 MetTrpAspValPheCysTrpArgMetLysAspValAsnCysThr***LysValGlyTyr 161
 DB 184 ATGGGAGATGCTTCTGCTATCGGATGAAGATGAAGTGAACCTGCACCTGCAGGTGGGCTAT 243
 QY 162 ValGlyAspGlyPheSerTyrSerGlyAsnLeuLeuGlnValLeuMetSerPheProSer 181
 DB 244 GTGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCTGTATGCTTCCCTCA 303
 QY 182 LeuThrAsnPheLeuThrGluValLeuAlaIleTyrSerAsnSerSerAlaArgGlyArgAla 201
 DB 304 CTCACAACTTCTTGACGGAAGTGTGGCTTATTCACACAGCTCAGCTCGAGGGCGTGCA 363
 QY 202 PheLeuGluHisLeuThrAspLysSerIleArgGlyThrLeuPheValProGlnAsnSer 221
 DB 364 TTTCTAGAACACCTGACTGACTGCTCCATCCGCGGACCCCTCTTTGTGCCACAGACAGT 423
 QY 222 GlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisHisLeuAlaAsnVal 241
 DB 424 GGGCTGGGGGAGATGAGACCTTGTCTGGGGGGGACATCGAGACACACCTGCCAATGTC 483
 QY 242 SerMetPhePheTyrAsnAspLeuValAsnGlyThr-ThrLeuGlnThrArgLeuGlySer 261
 DB 484 AGCATGTTTTTCTACATGACCTTGTCAATGGCCACCCNACCTCGCAACGAGGGTGGGAAG 543
 QY 261 rLysLeuLeuIleThrAspArgGlnAspProLeu-HisProThrGluThrArgCysVal 280
 DB 544 CAAGCTGCTCATCACTGCCAGCCAGGCCACCCACTNCCNACACGAGGAGACCGAGTTGTT 603
 QY 281 AspGlyArgAspThrLeuGluTrpAspIleCysAlaSerAsnGlyIleThrHisValIle 300
 DB 604 GATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCTCCATGGATCATTCATTCATTCATT 663

Qy 200 gAlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAs 220
 Db 1801 TGCATTTCTAGAACACCTGACTGACCTGCCATCGCGGACCCCTCTTTGTGCCACAGAA 1860
 Qy 220 nSerGlyLeuGlyGluAanGluThrLeuSerGlyArgAspIleGluHisLeuAlaAs 240
 Db 1861 CAGTGGGCTGGGGAGAAATGAGACCTTCTGTGGCGGACATCGAACACCCCTCGCCAA 1920
 Qy 240 nValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuG1 260
 Db 1921 TGTGACGATGTTTTTCTACAATGACCTTGTCATGTCACACCCCTGCCAACAGAGGCTGGG 1980
 Qy 260 ySerIysLeuLeuIleThrAspArgGlnAsp 270
 Db 1981 AAGCAAGTGTCTCATCTACTGCCAGCAGGAC 2011
 RESULT 13
 ABL90758
 ID ABL90758 standard; cDNA; 3681 BP.
 XX AC ABL90758;
 XX DT 24-MAY-2002 (first entry)
 XX DE Human polynucleotide SEQ ID NO 1320.
 XX KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
 KW vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein; gene; ss.
 XX OS Homo sapiens.
 XX SS WO200190304-A2.
 XX PN 29-NOV-2001.
 XX PD 18-MAY-2001; 2001WO-US016450.
 XX PF 19-MAY-2000; 2000US-0205515P.
 XX PR (HUMA-) HUMAN GENOME SCI INC.
 XX PA Birse CE, Rosen CA;
 XX PI WPI; 2002-122018/16.
 XX DR P-PSDB; ABB90349.
 XX PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
 PT prevention of neural, immune system, muscular, reproductive,
 PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
 PT disorders.
 XX PS Claim 4; SEQ ID NO 1320; 208lpp + Sequence Listing; English.
 XX CC The invention relates to novel genes (ABL9449-ABL90853) and proteins
 CC (ABB9040-ABB9044) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
 CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
 CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
 CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
 CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing
 CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
 CC infectious diseases such as viral, bacterial, fungal and parasitic
 CC infections. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 3681 BP; 648 A; 1172 C; 1123 G; 731 T; 0 U; 7 Other;
 SQ Alignment Scores:
 Pred. No.: 2,79e-65 Length: 3681
 Score: 661.50 Matches: 145
 Percent Similarity: 53.44% Conservatives: 49
 Best Local Similarity: 39.94% Mismatches: 148
 Query Match: 35.47% Indels: 21
 DB: 6 Gaps: 7
 US-09-466-778B-11 (1-353) x ABL90758 (1-3681)
 Qy 2 ThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeuAsn 21
 Db 2235 ACCGGCTGAACACACGCGCTGTGAGTCCACGACGGCTACGTAGCGATGAGTGCAG 2294
 Qy 22 Cys--GluProGluGlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis 40
 Db 2295 TGTCTGGAGGAGTCGGAACACACCTGTGGACCGCTTGGGCCAGCCGCCCTTGCAC 2354
 Qy 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60
 Db 2355 TCAGATGCCATGTGCACTGACCTGCACTTCCAGGAGAACGGGTGGGTTTTCACCTC 2414
 Qy 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80
 Db 2415 CAGGCCACACGCGCCCTTATGCTGCTGAACCTTTCGAGGCTGAGCGGATGCGAAGCA 2474
 Qy 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr**GlnLysAlaLysTyrHis 100
 Db 2475 CAGGAGCGCTGCTTGTCTTCTCCCTCAGCTCTCTGCTGCCAGCAGCTGGGCTTCCAC 2534
 Qy 101 LeuCysSerAlaGlyTyrLeuGluThrGlyValAlaTyrProThrAlaPheAlaSer 120
 Db 2535 CTGTGCTCATGGCTGGCTGGCCAAATGGCTCCACTGCCACCTGTGGTTTCCCTGTG 2594
 Qy 121 GlnAsnCysGlySerGlyValValGlyIleValAspTyrGlyProArgProAsnLysSer 140
 Db 2595 CGGACTGTGGCAATGTGCGGGTGGGCATAGTCAGCCTGGGTGGCCGCAAGAACCTCTCA 2654
 Qy 141 GluMetTrpAspValPheCysTyrArgMetLysAspValAsnCysThr**LysValGly 160
 Db 2655 GAACGTGGGATGCTCTTCCGTGTCAAGATGTGGCTGCCATGCCAATGGC 2714
 Qy 161 TyrValGlyAspGlyPheSer---TyrSerGlyAsnLeuGlnValLeuMetSerPhe 179
 Db 2715 TTCGTGGGTGAGCGGATCAGCAGCTGCAATGGGAAGCTGTGGATGTGTGGCTGCCACT 2774
 Qy 180 ProSerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGly 199
 Db 2775 GCCAACTTCTCCACTTCTATGGGATGCTATTGGGCTATGCCAATGCCACCGAGGGGT 2834
 Qy 200 ArgAlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGln 219
 Db 2835 CTCGACTTCTGGACTTCTGGATGAGTACGTACGTATAGACACACTTCTGCTCCCTGTC 2894
 Qy 220 AsnSerGlyLeuGlyGluAanGluThrLeuSerGlyArgAspIleGluHisLeuAla 239
 Db 2895 AATGAAGCTTTGTGGCAACATGACGCTGAGTGGCCAGACTTGGAGTGCATGCTCC 2954
 Qy 240 AsnValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeu 259
 Db 2955 AACGCCACCTCTCTAAGTGGCAAC---GCCAGCCAGGGGAAGTTGCTTCCGCCCTCA 3011
 Qy 260 GlySerLysLeuLeuIleThrAspArg-----GlnAspProLeuHisProThrGlu 276
 Db 3012 GGCCTGAGCTTCATCATCAGTGCAGCGAGGCCCTGACAAACAGTTCCTGGGCCCTTGTGCC 3071
 Qy 277 ThrArgCysValAspArgGlyArgAspThrLeuGluThrPheAspIleCysAlaSerAsnGlyLe 296
 Db 3072 CCAGGACAGTTGTGGTTAGCGGTATCATTTGTGGGACATCATGCGCTTCAATGGCATC 3131


```
QY 180 ProSerLeuThrAsnPheLeuThrGluValLeuAlaTySerAsnSerSerAlaArgGly 199
Db   ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
688 GCCAACTTCTCCACCTTCTATGGGATGCTATTGGGCTATGCCAATGCCACCCAGCGGGT 747
QY 200 ArgAlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGln 219
Db   ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
748 CTCGACTTCTCGGACTTCTGGATGATGAGCTCACGTATAGACACTCTTCGTCCTGTC 807
QY 220 AsnSerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisLeuAla 239
Db   ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
808 AATGAAGGCTTTGTGGACACACATGAGCTGAGTGGGCCCAACTTGGAGCTGCATGCTCC 867
QY 240 AsnValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeu 259
Db   ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
868 AACGCCACCTCTCAAGTGCCAAAC--GCCAGCCAGGGGAAGTTGCTTCGGGCCCACTCA 924
QY 260 GlySerLysLeuLeuIleThrAspArg-----GlnAspProLeuHisProThrGlu 276
Db   ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
925 GGCTCAGCCTCATCATCATCATGAGTACGAGGCCCTGACAAACAGTTCTGGGCCCTGTGGCC 984
QY 277 ThrArgCysValAspGlyArgAspThrLeuGluTyrAspIleCysAlaSerAsnGlyIle 296
Db   ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
985 CCAGGGACAGTTGTTAGCCCGTATCATTTGTGTGGGACATCATGGCCTTCATATGGCATC 1044
QY 297 ThrHisValIleSerArg***LeuLysAlaProProAlaProValThrLeu----- 313
Db   ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
1045 ATCCATGCTCTGGCCAGCCCTCTCTGGCACCCGCCACAGCCCGGAGCTGTGGGCGCNT 1104
QY 314 -----**HisThrGlyLeuGly**GlyIlePhe*****IleIleLeuVal 329
Db   ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
1105 GAAGCCCCACCTGTGGCGGAGCGGTGGGGCT-----GTGCTTGCC 1146
QY 330 ThrGlyAlaVal-----AlaLeuAlaAlaTySerTyrPheArgIleAsnArgLysThr 347
Db   ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
1147 GCTGGAGCACTGTGTGGCTTGTGTGGCGGAGCTCTTACCTCGTCCCGGAGCAAGCCC 1206
QY 348 IleGlyPhe 350
Db   ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
1207 ATGGGCTTT 1215
```

Search completed: April 1, 2004, 12:37:58
Job time : 586 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 1, 2004, 09:56:50 ; Search time 4925 Seconds
(without alignments)

3106.617 Million cell updates/sec

Title: US-09-466-778B-11

Perfect score: 1865

Sequence: 1 MTGPKHCKCKSHYVGDGL.....ALAAYSYFRINRKTIGFXHF 353

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-O=/cgn2_1/USPTO.spool/US09466778/runat_01042004_095642_10579/app_query.fasta_1.519
-DB=GenEmbl -PMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09466778@cgn1_1_3731@runat_01042004_095642_10579 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEX=7
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEX=7

Database :

GenEmbl:

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_ov:*
- 5: gb_pat:*
- 6: gb_ph:*
- 7: gb_pl:*
- 8: gb_pr:*
- 9: gb_ro:*
- 10: gb_sts:*
- 11: gb_sy:*
- 12: gb_un:*
- 13: gb_vl:*
- 14: gb_vl:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_or:*
- 21: em_ov:*
- 22: em_pat:*
- 23: em_ph:*
- 24: em_pl:*
- 25: em_ro:*
- 26: em_sts:*
- 27: em_un:*
- 28: em_un:*

- 29: em_vi:*
- 30: em_htg_hum:*
- 31: em_htg_inv:*
- 32: em_htg_other:*
- 33: em_htg_mus:*
- 34: em_htg_pin:*
- 35: em_htg_rnd:*
- 36: em_htg_mam:*
- 37: em_htg_vrt:*
- 38: em_sy:*
- 39: em_htgo_hum:*
- 40: em_htgo_mus:*
- 41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Length	DB ID	Description
1	1772	95.0	4492	9	AY227444 Homo sapi
2	1772	95.0	4575	9	AK024503 Homo sapi
3	1772	95.0	5604	9	AK074051 Homo sapi
4	1772	95.0	8214	9	AB052958 Homo sapi
5	1772	95.0	8251	9	AY311388 Homo sapi
6	1772	95.0	8266	9	AJ295695 Homo sapi
7	1769	94.9	3260	6	AX686401 Sequence
8	1769	94.9	3260	6	AX686401 Sequence
9	1558	83.5	3625	6	AX149461 Sequence
10	1558	83.5	3642	9	AL133021 Homo sapi
11	1558	83.5	8495	6	AX686192 Sequence
12	1554	83.3	8444	6	AX686192 Sequence
13	1514	81.2	4706	10	AY007370 Rattus no
14	1487	79.7	8157	10	AF364951 Mus muscu
15	1203	64.5	2011	6	AX149447 Sequence
16	664.5	35.6	3911	10	BC031166 Mus muscu
17	664.5	35.6	7931	10	AF290914 Mus muscu
18	661.5	35.5	7909	9	D87433 Homo sapien
19	661.5	35.5	7916	9	AB052956 Homo sapi
20	652.5	35.0	7870	9	HS275213 Homo sapi
21	643	34.5	1500	4	ETA50060 Bos tauru
22	642	34.4	2160	6	AX748299 Sequence
23	642	34.4	2160	9	AK093774 Homo sapi
24	632.5	33.9	1482	6	AR269667 Sequence
25	626	33.6	1804	6	AX149449 Sequence
26	581	31.2	2863	6	AX747085 Sequence
27	581	31.2	2863	9	AK091600 Homo sapi
28	535	28.7	2483	6	AX149459 Sequence
29	535	28.7	2512	6	AX149455 Sequence
30	479.5	25.7	249541	2	AC095672 Rattus no
31	479.5	25.7	249774	2	AC136196 Rattus no
32	464	24.9	199310	9	AC112215 Homo sapi
33	464	24.9	217029	9	AC105934 Homo sapi
34	458.5	24.6	146283	2	AC022277 Homo sapi
35	458.5	24.6	180064	9	AC025265 Homo sapi
36	458.5	24.6	180525	2	AC063946 Homo sapi
37	458.5	24.6	184435	2	AC141618 Homo sapi
38	458.5	24.6	238220	2	AC107727 Mus muscu
39	414	22.2	1249	10	BC022136 Mus muscu
40	400.5	21.5	202968	2	AC137422 Rattus no
41	388.5	20.8	204736	2	AC113310 Mus muscu
42	388.5	20.8	247224	2	AC112790 Mus muscu
43	388.5	20.8	343587	2	AC113471 Mus muscu
44	388.5	18.7	6824	6	AX781121 Sequence
45	267.5	14.3	769	9	HUMZD93D04

ALIGNMENTS

RESULT 1

AY227444 4492 bp mRNA linear PRI 23-APR-2003
LOCUS Homo sapiens hyaluronan receptor for endocytosis precursor (HARE)
DEFINITION mRNA, partial cds.
ACCESSION AY227444
VERSION AY227444.1 GI:28848629
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 4492)
AUTHORS Zhou, B., McGary, C.T., Weigel, J.A., Saxena, A. and Weigel, P.H.
TITLE Purification and molecular identification of the human hyaluronan
receptor for endocytosis
JOURNAL Glycobiology 13 (5), 339-349 (2003)
MEDLINE 22588541
PUBMED 12626425
REFERENCE 2 (bases 1 to 4492)
AUTHORS Weigel, P.H.
TITLE Direct Submission
JOURNAL Submitted (29-JAN-2003) Biochemistry and Molecular Biology,
University of Oklahoma Health Sciences Center, BMSB Room 860,
Oklahoma City, OK 73104, USA
FEATURES
source
1..4492
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
<1..4492
/gene="HARE"
<1..4251
/gene="HARE"
/codon_start=1
/product="hyaluronan receptor for endocytosis precursor"
/protein_id="AAO39681.1"
/db_xref="GI:28848630"
/translation="SLPNLLMRLEQMPDYSIFRGVYIIQYNLANAIEADAVTVFAPNN
NAIENYIREKVLSEEDVLVHYVLEBKLLKNDLHNGMHRETMLGFYSYFLSLFND
QLYNEAPINTVATDKGVHIGLVKLEIKNRCDNDNTIILGRCTCSSELTCTPP
NVCFNGICLVGVNGTGCCEGEGFSGTACETGKGIHCDQACSVHRCNQGPI
GDGSCDVWRGVKCDNATNDNCGTSCANTLNSDGTASCACAGPQNGGTC
TAINACEISNGGCSAKAKRTTPGRVCTCKAGTGDGIVCLEINPLENHGCDKN
AECTGTGPNQACNCLPAYTGDGCTLIYVCLTKNGGCSFALCNHTGQVETCTCK
PNYIGDFTCRGSIYELPKPKTSQVFEQLOEHFVKDLVGPBPVTFAPLSAADFDE
ARVNDWKYGLMPQVLRHYVACHQLLENLKLISNATSLQGEFIVSVSSTYINN
KAKLISDIISTNGIIVHIIIDKLLSPNLLITPKNSGRILQNLITLATNNGYIIRFSNL
TODGLLSVIITDPIHTVTLFWPTDQALHALPABQQDFLFDQNDKDLKEYLKFVIR
DAKVLADPLTSATWKTQSGSELSVKCGAGRIDGLFLNGQTCTRIVORELLFDLGVAY
GIDCLLIDPTLGGRCDTFTPDASGEGSCVNTSPCRWSPKGVKQKLYNLFPKRN
LEGCFRCSLVITQIPRCCKGVGRDQACPGDPAPCNRRGVCLDQXSATGECKNTG
PNTGACEMCWEPGPGDCLPGCSDHGQCDGIGTSGQCLCTGWTGSPCSTQAVLPA
VCTPFSAHATCKENNTCEKNDIEGDTITCTVVDQKNGGCAKVARCSQKGTKVS
CSQKGYKGDHSCSTEDIPCADLNGGCHHEATKMTGPGHKCKSHYVYGDGANCE
PEQPIFDKQSGCHADAKCVLIDFQDTTVGVPLHRSPLGQYKLFTHDKAREACANE
AATMATYNQLSVAQAKYKLSAGWLETGRVAYPTAFASQNCGGVGVIVDYGPRPNK
SEMDVFCYRMKDVNCTCKGVGDPGSCGNLLQVLMSPSLNFTLTVLAYNSSSA
RGRAPLEHLTDLISRGTLFVQNSGLGENETLSGDIIEHLNYSMEFYNDLVNGLTIL
QTRLSKULLITFAISQDPLQPTETRFVDGRLAILQWIDFASNGIHHVIRSLKAPPAPVIL
THTGLIGAFITAILVTGVALAAYSYFRINRRTIGFQHFSEEDINVAALGKQOOPEN
ISNLYESTTSAPPEPSYDPFTDSERQLEGNDFLRTL"
1..4248
mat_peptide
/gene="HARE"
/product="190 kDa hyaluronan receptor for endocytosis"
/function="mediates endocytosis of hyaluronan or
chondroitin sulfate via the clathrin coated pit pathway"
/note="generated by proteolytic processing of a larger
protein; probably a product of the HARE/Stablin-2 gene;
190 kDa protein purified from spleen"

ORIGIN

Alignment Scores:

Pred. No.: 2,61e-164 Length: 4492
Score: 1772.00 Matches: 334
Percent Similarity: 95.18% Conservative: 2
Best Local Similarity: 94.62% Mismatches: 17
Query Match: 95.01% Indels: 0
DB: 9 Gaps: 0
US-09-466-778B-11 (1-353) x AY227444 (1-4492)
QY 1 MetThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeu 20
Db 3022 ATGACAGCCCGCGGCAAGCAAGTGTGAGTGTAAAGTCACTATGTGCGAGATGGGCTG 3081
QY 21 AsnCysGluProGluGlnLeuProIleAspArgCysLeuGlnAspAenGlyGlnCysHis 40
Db 3082 AACTGTGAGCGGAGAGAGCTGCCCATTTGACCGCTGTACAGACATGGCAGTGCAT 3141
QY 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60
Db 3142 GCAGACGCCAAATGTGTGACCTCCACTTCAGGATACCACTGTGGGTGTTCATCTA 3201
QY 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80
Db 3202 CGCTCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGGCTGTGCCAAC 3261
QY 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis 100
Db 3262 GAAGCTGGACCATGGCAACCTACACAGCTCTCTATGCCCAAGAGGCCAAGTACCAC 3321
QY 101 LeuCysSerAlaGlyTyrLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120
Db 3322 CTGTGCTCAGCAGGCTGGTGGAGACCGGGGGTGTGCTACCCACACAGCCTTCGCCTC 3381
QY 121 GlnAsnCysGlySerGlyValValGlyLeuValAspTyrGlyProArgProAsnLysSer 140
Db 3382 CAGAACTGTGCTGTGTGTGGTAGTGGACTATGAGCTATGAGCTAGACCCACACAGAT 3441
QY 141 GluMetTyrAspValPheCysTyrArgMetLysAspValAenCysThr***LysValGly 160
Db 3442 GAATGTGGAGTGTCTTCTGCTATCGATGAAAGATGTGAATGCACCTGCAAGGTGGC 3501
QY 161 TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuLeuGlnValLeuMetSerPhePro 180
Db 3502 TATGTGGAGATGGCTTCTCATGAGTGGGAACTGTGTGAGGTCTCTGATCTCTCC 3561
QY 181 SerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArg 200
Db 3562 TCCTCACAACCTTCTTCAGCAAGTGTGGCTTATTCACAACAGCTCAGCTCGAGGCGGT 3621
QY 201 AlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsn 220
Db 3622 GCATTTCTAGAACCTTGACTGACCTGTCCATCCGCGGCACCCCTCTTTGTGCCACAGA 3681
QY 221 SerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisLeuAlaAsn 240
Db 3682 AGTGGCTGGGGAGATGAGACCTTCTGTGGCGGGACATCGAGACCACTCGCAAT 3741
QY 241 ValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGly 260
Db 3742 GTCAGCATGTTTTCTCAATGACCTTGTCAATGGCACCCCTCGCAACAGAGCTGGGA 3801
QY 261 SerLysLeuLeuIleThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal 280
Db 3802 AGCAAGCTGCTCATCTGCGCAGCCAGGCCCTCCCAACCGAGGAGACCAAGTTTGT 3861
QY 281 AspGlyArgAspThrLeuGluThrAspIleCysAsnSerAsnGlyIleThrHisValle 300
Db 3862 GATGGAAGAGCATTTCTGCAATGGGACATCTTTGCCCTCAATGGGATCATTCATGTCA 3921
QY 301 SerArg***LeuLysAlaProAlaProAlaProValThrLeu***HisThrGlyLeuGly*** 320
Db 3922 TCCAGGCGCTTTAAAGCACCCTGCCCCCGTGTACCTTGACCCACACTGGCTTGGGAGCA 3981

Db 3795 GATGGAAGAGCAATCTCGAGTGGGACATCTTGGCTCCATGGGATCAATTCATGTCATT 3854

Qy 301 SerArg***LeuLysAlaProProAlaProValThrLeu***HisThrGlyLeuGly*** 320

Db 3855 TCCAGGCGCTTTAAAGACACCCCTCCCGCGTGACCTTTGACCCACACTGGCTGGGAGCA 3914

Qy 321 GlyIlePhe*****IleIleLeuValThrGlyAlaValAlaLeuAlaAlaTyrSerTyr 340

Db 3915 GGGATCTCTTTGGCATCATCTCGTGAGTGGGGCTGTGCTTGGCTGCTTACTCTTAC 3974

Qy 341 PheArgIleAsnAtgLyThrIleGlyPhe***HisPhe 353

Db 3975 TTTCCGATTAACCGAGACATCGCTTCCAGCATTTT 4013

RESULT 3

AK074051

LOCUS Homo sapiens mRNA for FLJ00122 protein. linear PRI 15-FEB-2002

DEFINITION AK074051

ACCESSION AK074051.1 GI:18676449

VERSION fib (full insert sequence).

KEYWORDS Homo sapiens (human)

SOURCE

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS Jikuya, H., Takano, J., Nomura, N., Kikuno, R., Nagase, T. and Ohara, O.

TITLE The nucleotide sequence of a long cDNA clone isolated from human spleen

JOURNAL Published Only in Database (2002)

REFERENCE 2 (bases 1 to 5604)

AUTHORS Jikuya, H., Takano, J., Nomura, N., Kikuno, R., Nagase, T. and Ohara, O.

TITLE Direct Submission

JOURNAL Submitted (21-JAN-2002) Takahiro Nagase, Kazusa DNA Research Institute, Department of Human Gene Research; 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: cdnainfo@kazusa.or.jp, URL: http://www.kazusa.or.jp/NEDO, Tel: 81-438-52-3913, Fax: 81-438-52-3914)

COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert and 5'- & 3'-end one pass sequencing; Research Association for Biotechnology; cDNA library construction and clone selection; Kazusa DNA Research Institute.

FEATURES

source

1..5604

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="FLJ00122"

/tissue_type="spleen"

/dev_stages="adult"

/note="vector:pBluescriptII SK plus"

gene

1..5604

/gene="FLJ00122"

<1..5213

/gene="FLJ00122"

/note="Start codon is not identified."

/codon_start=3

/evidence=not experimental

/product="FLJ00122 protein"

/protein_id="BAB04877.1"

/db_xref="GI:18676450"

/translation="DGSAGRLCDKQTSACGPYVQFCHHATCEYNSGTASGICAKGYE

GDGTLCEMDPCTGLTPGGCSRNAECIKGTGTHTCVQCGWTNGNDCSEINNCLLP

SAGHCNDSALCYVPGQNECECKKFRNGIDCEPITSLEQTKGCHPLASQOSTSS

GVMSVCEQEGYEGDFCLCYNAVALSLSEAAIFNFWINNASSLOPILSATSNLTIVL

PSQAATDMQDDEKSWLSQSNIPALIKYHMLCTYEVADLQTLSSDMLATSLQGNF

LHLAKVQNTIETGASVDGNAATNGVHIINKVLPQRRLTSGSLNMLMRLEQFDE

YSIFRGYIQYNLANAEADAYTVFAPNNANENYIREKKVLSLEEDLRYHVHVLDE

KLLKNDLHNGHMETLGFYSFLFFLHNDQLYVNEAPINYNTVATDKVHGLGKYL

EIQKRDNDNDTTIIRGCRITCSSELTCPFGTKSLGNEKRRCIYTSYFMGRRTIFIGC

QPKQRTVITVRECCAGFPQCPGNAQVCFNGICLDGVNGTGVCECGEFSGT

ACETCTEGYGIHCDQACSCVHGRCNQGPLDGSDCDVGWRGVHCDNATTEDCNCT

ORIGIN

Alignment Scores:

Pred. No.: 3,46e-164 Length: 5604

Score: 1772.00 Matches: 334

Percent Similarity: 95.18% Conservative: 2

Best Local Similarity: 94.62% Mismatches: 17

Query Match: 95.01% Indels: 0

DB: 9 Gaps: 0

US-09-466-778B-11 (1-353) x AK074051 (1-5604)

Qy 1 MetThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeu 20

Db 3984 ATGACAGCCCGGGCAGACAGTGTGAGTGTATAAGTCACTATGTCGAGATGGGCTG 4043

Qy 21 AsnCysGluProGluInLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis 40

Db 4044 AACTGTGAGCGGAGCAGCTGCCCAITGACCGTGTCTACAGGACAAATGGCAGTGCAT 4103

Qy 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60

Db 4104 GCGACGCCAAATGTGTCGACCTCCACTCCAGATACCACTGTGGGGTGTTCACATCA 4163

Qy 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80

Db 4164 CGCTCCCACTGGGCGAGTATAAGTGAACCTTTGACAAAGCCAGAGAGGCTGTGCCAAC 4223

Qy 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis 100

Db 4224 GAAGCTGCGACCATGGCAACCTCAACACAGCTCTCTTATGCCAGAGGCAAGTATCCAC 4283

Qy 101 LeuCysSerAlaGlyTyrLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120

Db 4284 CTGTGCTCAGCAGCTGGCTGGAGACCGGGCGGTGGCTACCCACAGCTTCGCTCC 4343

Qy 121 GlnAsnCysGlySerGlyValValGlyIleValAspTyrGlyProArgProAsnLysSer 140

Db 4344 CAGAACTGTGGCTCTGTGTGTGGTATAGTGGACTATGGACTAGACCTAGACCAACAGAGT 4403

Qy 141 GluMetTrpAspValPheCysTyrArgMetLysAspValAsnCysThr***LysValGly 160

Db 4404 GAAATGTGGGATGTCTTCTCTATCGGATGAAGAAGTGTGAACCTGCACCTGCAGAGGGGGC 4463

Qy 161 TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuLeuGlnValLeuMetSerPhePro 180

Db 4464 TATGTGGAGATGGCTTCTCATCGATGGGAACCTGCTGCAGGTCTCTGATGCTCTCC 4523

Qy 181 SerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArg 200

Db 4524 TCACCTACAAACTTCTCTGACGGAAGTGTGGCTATTCCAAACAGCTCAGCTCAGGCGGT 4583

Qy 201 AlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsn 220

Db	4584	GCATTCTAGAACACCTGACTGCTGCTCATCCGGGACCTCTTTGTGCACAGAAC	4643	
Qy	221	SerGlyLeuGlyGluAenGluThrLeuSerGlyArgAspIleGluHisHisLeuAlaAen	240	
Db	4644	AGTGGGCTGGGGAGAAATGAGACCTTGTCTGGGCGGACATCAGACACCTCGCAAT	4703	
Qy	241	ValSerMetPhePheTyrAsnAspLeuValAenGlyThrThrLeuGlnThrArgLeuGly	260	
Db	4704	GTGAGCATGTTTCTTACCAATGACCTTGTCAATGGCACCACCTCGCAACGAGGCTGGGA	4763	
Qy	261	SerLeuLeuLeuLeuThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal	280	
Db	4764	AGCAAGCTCTATCATCTCCAGCCAGACCACTCAACAGGAGACCAAGTTTGT	4823	
Qy	281	AspGlyArgAspThrLeuGluTrpAspIleCysAlaSerAsnGlyIleThrHisValIle	300	
Db	4824	GATGGAAGAGCCATCTGCGTGGGACATCTTGGCTCCATGGGATCATTCATGTCAT	4883	
Qy	301	SerArg***LeuLeuAlaProAlaProAlaProValThrLeu***HisThrGlyLeuGly	320	
Db	4884	TCCAGGCTTTAAAGACACCCCTGCCCCCGTGACCTTGACCCACACACTGGCTTGAGCA	4943	
Qy	321	GlyIlePhe*****IleIleLeuValThrGlyAlaValAlaLeuAlaAlaTyrSerTyr	340	
Db	4944	GGATCTCTTTGGCCATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	5003	
Qy	341	PheArgIleAsnArgGlyThrIleGlyPhe***HisPhe 353		
Db	5004	TTTCGGATAAACCGGAGAACATCGCTTCCAGCATTTT 5042		
RESULT 4				
LOCUS	AB052958	8214 bp	mRNA	linear
DEFINITION	Homo sapiens mRNA for FELE-2, complete cds.			
ACCESSION	AB052958			
VERSION	AB052958.1	GI:22779440		
KEYWORDS	Homo sapiens (human)			
SOURCE	Homo sapiens			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS	Adachi, H. and Tsujimoto, M.			
TITLE	FEEL-1, a novel scavenger receptor with in vitro bacteria-binding and angiogenesis-modulating activities			
JOURNAL	J. Biol. Chem. 277 (37), 34264-34270 (2002)			
MEDLINE	2206516			
PUBMED	12077138			
REFERENCE	2 (bases 1 to 8214)			
AUTHORS	Adachi, H.			
TITLE	Direct Submission			
JOURNAL	Submitted (25-DEC-2000) Hideki Adachi, RIKEN, Cellular Biochemistry Laboratory; Hirosewa 2-1, Wako-shi, Saitama 351-0198, Japan			
	(E-mail: adachi@postman.riken.go.jp, Tel: 81-48-467-9372, Fax: 81-48-462-4670)			
FEATURES	Location/Qualifiers			
source	1..8214			
	/organism="Homo sapiens"			
	/mol_type="mRNA"			
gene	/db_xref="taxon:9606"			
	1..8214			
	/gene="FELE-2"			
CDS	183..7838			
	/gene="FELE-2"			
	/note="fasciclin, EGF and laminin type EGF containing scavenger receptor-2"			
	/codon_start=1			
	/product="FELE-2"			
	/protein_id="BAC15608.1"			
	/db_xref="GI:22779441"			
	/translation="MMQLVIFCLGLVQNFCSPAFTTGGARRCDKSLTLTIRCR SCALNLGVKPDGYMTITSSGVGRDCRYTFEVRTLSLPGCRHCRKDYLPQRCPP			

GRWPDCLIECPGAGSPNGRSCAEGMEGNTCSQEGGCTGACTACADDDLPSPSC
SSVNCVHGVCNSGLDGTCEYAYTGPCKDKPIPECAALLCPENSRCSPSTBDEN
KLECKLNYRGDKYCDIPNCLRKI CHPHACTYLGPNRHSCTCQEGYRGDQVCL
PVDQINFGRCPTKSTYCKVDPQSCHECKERYQNFVPGVGCSDMTICKSDNPHR
NACTTAVPGRTECIQKQYVGDGLTCYGNMERLELNRELNREPRKQWGLTFTI SILD
KAYAWPLSKLGFVTLPTDQKGLFNVELVDNKAAYFVKLHI IAGQNNIEYNN
TDMFTLLTKSGEIEFNSDKQIKLKHGKKVKIKQDI IASNGLLHILDRAMDKL
BHTFENNEOTIMTLQPRYSKRSLEBETNHALDDEGCGPVTI FVPNNALNNM
KQTDUDYLLSPGSKLLELVRYHIVPTOLEVATLI STPHRSNANQLI QNTTNDG
OILANDVAMELEITAKNGRYTLTGVLIPSVIFILPHRCDETRKMLGTCVCSCL
YIWSRCPANSPTALTTHRCVYSGFGLSKGCYCNATVKI PKCGYKFGPDNCQC
PGFSPNCSNGNQCADSLGNGTCICEGFGSQCFQCPDNKYPGRNCKGLCVHGT
CNRITDSGACLTGTCRDSAGRLCDKTSAGPYVQVCHIQATCEYNGTASCCTKA
GYEGDGLTCSMDPCTGLTGGCSRNAECIKGTGTHTCVQOHTGNGRDCSEINCA
LLPAGGCHDNASCLYVGPONECECKKFRNGIDCEPITSCLEQTKCHPLASCQS
TSSGVMSVCCEGVEGDFLYCGRNAVELSFLSEAAI FNRINNASLOTUSANLT
VIVPSQATEMDQDEKSFWSQSNIPALIKIHMLLGTYRVADLOTLSSDMLATSLQ
GNFLHLAKGNTITIEGASIVDGNAAATNGVHIINKVLVFORLTGSLPMLMLEQ
MPDYSIFGYIIQYNLANAIEADATVFPNNNAIENYIREKKVLSLEEDVLRHV
LEEKLKNDRHNGHRETMGLFSYFLPHNDOLYNEAPINVTNATDKGVHGLG
KVLEIKNDRNDNDTIIIRGRCTCSSLTCPTFGTKSLGNEKRCIVTSYPMGRITLF
IGCOPKCVTVITRECCAGFGPQCPGQPCFNAQNVCFNGICLDGNGTGVCECEGF
SGTACTCEBKYGIHCDQACSCVHRCNQGPDGSDCDVGRWGRHCDNATDENC
NGTCHTSANLUNSDGTASCKAAGFQNGTICTAINACEISNGSCSAKADKRTTPG
RRVCTKAGYTGDGIVCLEINPCLNHHGGCDNAECTQTPNAAACNCLPVTGDKV
CTLINVCITKNGGSEFAICNHTQVETCTCKPNIYIGDFTCRGSI VOELPKNPKTS
YFQLOHFHVKDLVGPFPVPAFLSAADFEARVKMDKYGKMPQVLRHVYVACHQ
LLENLKLIASNATLSOGEPIVISVQSIVYINAKKILSSDIISTNGVHIHIDKLLSP
KXLLITPKDNGRILQNTLTATNNGYIKFSNLIQDSGLLSVITDPIHTPTLFWPTD
QALHALFAEQDGLFNQDNKDKLEYLKFHVIRDAKVLADLPTSTAWKTQSGSELSV
KCGARDIGDLFNQDQTCRIVQRELLDVGAYIGDILLIDTLGRCDDTFTTDFDASG
ECGCVANTPSCPRMSKPKVKQKLYNLPKRNKGRCRSLVIOIPRCCQVFGFRD
CQACGPGDAPNNRGVCLDQYSATGCKNTGNTGNTACEMCPWEGPDCPLCGCSD
HGQCDGDTGSGOCLCETWTGSPCDTQAVLPVCTPPCSHAHCKENNTCECNDYE
GDGICTVDPKQDNGGCAKVARCKGTAKVSCQKGYKGDGSHCTEIDPCADGLN
GQCHEATCKMTGPKHCKECHKSHYVGDGLNCEPEQLPIDRLQDNGQCHAKADKVDL
HFQDITVGVFLRSPLQYKLTFDKAREACANEAAATWATYNQLSYAKVYHLCASGW
LETGRVAYPTAFASQNGCVGVIGVYDPRNKSEMDVFCYRMKDVNCTCKVGVVG
GFSCNGLDLOVMSFSLNFTLTVLAYSNSSAGRAFLBELHLDLSIRGLTFVQNSG
LGENETLSGRDIEHLNLSMFFYNDVNGTTLQRLGSKLLITASQDQPLQVTRFV
DGRALQWDIFASNGIILHSRPLKAPPVTLTHLGAGIFFAILLVGTGAVALAY
SYFRINRRTITGFQHFSEEDINVAALGKQKQPENISNPLYESTTSAPPEPSYDPTDSE
ERQLEGNDPLRLT"

ORIGIN

Alignment Scores:

Pred. No.: 5,6e-164 Length: 8214
Score: 1772.00 Matches: 334
Percent Similarity: 95.18% Conservatives: 2
Best Local Similarity: 94.62% Mismatches: 17
Query Match: 95.01% Indels: 0
DB: 9 Gaps: 0

US-09-466-778B-11 (1-353) x AB052958 (1-8214)

Qy 1 MetThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeu 20
Db 6509 ATGACAGCCCGGCAAGCACAAGTGTGAGTGTAAAGTCACTATGTGCGAGATGGGCTG 6568
Qy 21 AsnCysGluProGluGlnLeuProIleAspArgCysLeuGlnAspGlnGlyCysHis 40
Db 6669 AACTGTGAGCGGAGCAGCTGCCCAITTAGCCGCTGTACAGCAATGCCAGTGCAT 6728
Qy 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60
Db 6729 GCAGAGCCCAATGTGTGACCTCCACTTCCAGATACCACTGTGGGGTGTTCACATCTA 6788
Qy 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80
Db 6789 CGCTCCCACTGGGCGAGTATAGCTGACCTTTGACAGCCAGAGGCGCTGTGCCAAC 6848
Qy 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis 100

Db	6849	GAAGCTGCGACATGGCAACCTACAAACAGCTCTCTCTATGCCCGAGAGCCCAAGTACAC	6908
Qy	101	LeuCysserAlaslyTrpLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaser	120
Db	6909	CTGTGCTCAGCAGGCTGGTGGAGACCGGGGGTGGCTACCCACAGCTTCGCCCTC	6968
Qy	121	GlnAenCysGlySerGlyValValGlyIleValAlaPyrGlyProArgProAsnLysSer	140
Db	6969	CAGAACTGCGCTCTGGTGTGGTGGATAGTGGACTATGGACCTTAGACCCCAACAGAGT	7028
Qy	141	GluMetTrpAspValPheCysTyrArgMetLysAspValAsnCysThr***LysValGly	160
Db	7029	GAATGTGGAGTCTCTCTGCTATCGATGAAGATGGAATGCACTGCACGCTGGCGG	7088
Qy	161	TyrValGlyAspGlyPheSerTyrSerGlyAenLeuLeuGlnValLeuMetSerPhePro	180
Db	7089	TATGTGGAGATGGCTCTCATGCACTGGGAACCTGCTCAGGCTCTGATGCTCTCCOC	7148
Qy	181	SerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArg	200
Db	7149	TCACCTACAACTCTCTGACGAAGTGGGCTATTCACAGCTCAGCTCGAGGCGGT	7208
Qy	201	AlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAen	220
Db	7209	GCATTCTAGAACACCTGACTGACCTGCTCCATCCGCGGACCTCTTTGTGCCACAGA	7268
Qy	221	SerGlyLeuGlyGluAenGluThrLeuSerGlyArgAspIleGluHisLeuAlaAen	240
Db	7269	AGTGGGCTGGGGGAGAAATGAGACCTTGTCTGGCGGGGACATCGAGCACCATCCTC	7328
Qy	241	ValSerMetPhePheTyrAsnAspLeuValAenGlyThrThrLeuGlnThrArgLeuGly	260
Db	7329	GTCAGATGTTTTCTACAAATGACCTTGCTCAATGGCACCCCTGCAACAGAGGCTGGA	7388
Qy	261	SerLysLeuLeuIleThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal	280
Db	7389	AGCAAGCTGCTCATCTACCTCCAGCCAGGACCCACTCCACACGAGACCAAGTTGTT	7448
Qy	281	AspGlyArgAspThrLeuGluThrIlePheCysAlaSerAsnGlyIleThrHisValle	300
Db	7449	GATGAAGAGCCATTCTCGAGTGGGACATCTTGTCTCCATGGGATCATTCATGCTATT	7508
Qy	301	SerArg***LeuLysAlaProAlaProAlaProValThrLeu***HisThrGlyLeuGly	320
Db	7509	TCCAGGCTTTAAAGCACCCCTGCCCCGCTGACCTTGACCCACACTGCTGGGAGCA	7568
Qy	321	GlyIlePhe*****IleIleLeuValThrGlyAlaValAlaLeuAlaAlaTyrSerTyr	340
Db	7569	GGGATCTCTTTGCCATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTAC	7628
Qy	341	PheArgIleAsnArgLysThrIleGlyPhe***HisPhe 353	
Db	7629	TTTCGGATAAACCGGAGAACATCGGCTTCCAGCATTT 7667	
RESULT	5		
LOCUS	AY311388		
DEFINITION	Homo sapiens FEX2 mRNA, complete cds.		
ACCESSION	AY311388		
VERSION	AY311388.1	GI:32351284	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 8251)		
JOURNAL	Park,S.-Y. and Kim,I.-S.		
AUTHORS	FEX2, a novel cell adhesion molecule of Fas-1 superfamily mediates		
TITLE	cell-cell interaction		
REFERENCE	Unpublished		
AUTHORS	2 (bases 1 to 8251)		
TITLE	Park,S.-Y. and Kim,I.-S.		
	Direct Submission		

JOURNAL	Submitted (31-MAY-2003) Department of Biochemistry, School of Medicine, Kyungpook National University, 101 Dongin-dong, Jung-gu, Daegu 700-422, Korea
FEATURES	Location/Qualifiers
source	1. 8251
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/chromosome="12"
	/map="12q23.3"
CDS	205..7860
	/note="FAS1, BGF-like and X-link domain containing adhesion molecule-2"
	/codon_start=1
	/product="FEX2"
	/protein_id="AAP74958.1"
	/db_xref="GI:32351285"
	/translation="MMLQHLVIFCLGLVQVFCSPAETTGQARRCDKSLITIRTECR
	SCALNLGKCPDGYTMTSSGVGVRDCRYTEVRYTSLSLPGCHRICRKYLRCCP
	GRWPDCLCEPGGAGSPCNGRGSCABGMEGNGTSCQEGFGGTACETCADDNLPEPSC
	SSVNCVHGVCNSGLDGDGTCYSAVTGPKDKPIPECAALLCPENSRCSPTFEDSC
	KLECKLPNRYGDKYCDIPNPKIKIHPHAHCTYLPNHRHSCCTCQGYRGDQVCL
	PVDPQCNFGNCPKSTVKYDGPQSCHECKHYONFVPGVGSMTDICKSDNPCHR
	NACTTVAAGRTEGICQKGYVGDGLTCYGNIMERLRELTPRGKQGLTSLFSLLD
	KAYAPLSKGLPFTVLAPTDGLAGFNVELLDNKAQYFVKHLILAGOMNEMNN
	TDMYTITGSGEIPNSDKONQIKLKHGKKVKITQGDIIASNLHLIDRAMDKL
	EPTFESNNEOTIMTLOPRYSKFRSLIEETNLHALDEGVPYITFVPNEALNNM
	KDGLTYLLSPGKLLLELYRHYVPTFTEVATLSTPHIRMANQILQNTDNG
	QILANDAMERIEITAKNGRIYTLTLPSPVILPHRDETKRKLGTCTVCSCL
	YVMRCANSEPTALTFRCSVSGRFGSLKSGCARYCNATYIKPKCKGVPQNCOC
	PGFSNCSNGCADSLGGNGTICICREFGOGSOGFCSDPNKYPNKKCLCVHGT
	CNNRLDSCALTGTCDRDSAGRLCDKQTSACGYVQFCHHATCEVNSGTASICKA
	GYEGDGLCSMDPCTGLTPGGCNAECIKITGTHTCVCGQWNGRCDSEINNCC
	LLPSAGGCHNASCLYVPGQNECECKGFRNGIDCEPITSCLEQTKCHPLASQCS
	TSSGVWCVCQEGYEGDFLCYGAVALSFLSBAAPNRWINNASIQTLTSAVSNLT
	VLVPSQATDMDODEKSWLSQSNIPALI KYHMLGTYRVADLQTLSSDMLATSLQ
	GNFLHLAKVDNITIEGASIVDGNATNGVIHILNKVLPQRLTSLPMLLRLEQ
	MDYSIFGYIIQYNLANAIEADATVFAFNNAENIYIREKVLSEEDLVRYHV
	LEELKNDLHNGMRETMFLGSFYFLFNDQLYVNEAPINTVATDGVHGLG
	KVLBIQNRKNDNTTIRGRCTCSSELTCPFGTKSLGNEKRCIYTSYFMRITLF
	IGCPKCVRTITRECCAGFPQCPQCPQNAQVFCNGICLDGVNGTGCCEGEGF
	SGTACETCKGYIHCDQACSCHVRCNQPLGDGSCDGVWRGVCDDNATTEDC
	NRVCTSNACLTNSDGTASCKAFQNGTICTAINACEISNGCSACADCKRTTPG
	BRVCTKAGTYGDIIVCLEINPCLENGGCDKNAECTGTPNQACNCLPAYTGGKV
	CTLNVCLITKNGGCSBFAICNHTQVETCTCKENYIGDFTKGSITQELPKPKTS
	QYFQLQHFHVDLVGPFVTFAPLSAFAFDEARVDMWDKGLMPQVLRVHVACHQ
	LLLENKLI SNATLSQEPYIVISQSTYINNAKIISSDIIISNGIVHVIDLLSP
	KNLITPKNSGRILQNLTLTANNYIKFSLNLIQDSGLLSVTDPTHTPTLFWPST
	QALHALPABQDFLPNODNKKLKYLFKHYIRDAKVLAVLPTSTAWKTQGSLSV
	KCGAGRDGLFLNGQCTRIVORELLPDLGVAYGIDCLLIDPTLGGRCDTFTTDSG
	ECGSCVTPSPCRMSKPKYQKCLYNLPKFNLEGCEERCSLVIQIPRCCKGYFGRD
	COACPGGDFPCNNRGVCLDQYSATGCKNTGNGTACENMGWGRGPDCLPCGSD
	HGQCDGTITGSGQCLCTGWTGSPCDDTQAVLPVCTPPCSAHATCKENNTCECLDY
	GGCHEATCKMTPGKKECKSHVGDGLNCEPEALPIDRCLDQCHADAKCVDL
	HFQDTTGVHRLSPLOKLTPOKBEACENEAATMAYNQLSVAOKAKVHLCASGW
	LETGRVAPYAFASQNGSGVGVIVDYGPRNKSEMDVFCYRMKDNVCTCKGVYGD
	GFSCGNLLQVLMFSLNLTFLVAYNSSSARGAPLEHLTDLISIRGLTFLVFNQSG
	LGENETLSGRDIEHLANVSNFFYNDLVNGTTLQRLGSKLLITASQDLPQPTFRFV
	DGRAIQLWDIPASNGIITHVSRPLKAPPVLTHTGLGAGIFFAIILVTCGVALAAY
	SYFRINRTTIGFHFSEEDINVAALGKQKQENISNPLTYESTTSAPPEPSYDPTDSE
	ERQLEGNDPLRLT"

ORIGIN

Alignment Scores:			
Pred. No.:	5,63e-164	Length:	8251
Score:	1772.00	Matches:	334
Percent Similarity:	95.18%	Conservative:	2
Best Local Similarity:	94.62%	Mismatches:	17
Query Match:	95.01%	Indels:	0
DB:	9	Gaps:	0
	US-09-466-778B-11 (1-353) x AY311388 (1-8251)		

QY 1 MetThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeu 20
 Db 6631 ATGACAGCGCGGCAAGCAAGTGTGAGTGTAAAGTCACTATGTCGAGATGGCTG 6690

QY 21 AsnCysGluProGluGlnLeuProLleAspArgCysLeuGlnAspAsnGlyGlnCysHis 40
 Db 6691 AACTGTGAGCGGAGAGCTGCCATTGACCGCTGCTTACAGGACCAATGGCGAGTGGCAT 6750

QY 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60
 Db 6751 GCGAGCGCAAAATGTGCGACCTCCACTCCAGGATACCACTGTGGGGTGTTCATCTA 6810

QY 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80
 Db 6811 CGCTCCCACTGGCGCAGTATAAGCTGACCTTTGACAAAGCAGAGAGGCGCTGTGCCAAC 6870

QY 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis 100
 Db 6871 GAAGCTGCGACCATGGCAACCTCAACAGCTCTCTATGCCAGAGGCAAGTACCAC 6930

QY 101 LeuCysSerAlaGlyTyrLeuGluThrGlyValAlaTyrProThrAlaPheAlaSer 120
 Db 6931 CTGTGCTCAGAGGCTGGCTGGAGCGCGGGTGGCTTACCCACAGACTTCGCTCC 6990

QY 121 GlnAsnCysGlySerGlyValValGlyIleValAspTyrGlyProArgProAsnLysSer 140
 Db 6991 CAGAACTGTGGCTCTGGTGTGGTGGATAGTGGACTATGGACTAGACCAACAAGAGT 7050

QY 141 GluMetTyrAspValPheCysTyrArgMetLysAspValAsnCysThr***LysValGly 160
 Db 7051 GAAATGGGAGTGTCTCTCATGATGAAAGATGTGAACCTGCTGAGTCTGATGCTCTCCC 7110

QY 161 TyrValGlyAspGlyPheSerTyrSerGlyValAsnLeuGlnValLeuMetSerPhePro 180
 Db 7111 TATGTGGAGATGGCTCTCATGAGTGGAACTGCTGAGTCTGATGCTCTCTCCC 7170

QY 181 SerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerAlaArgGlyArg 200
 Db 7171 TCACCTCACAACTTCCTGACGGAAGTCTGGCTTATTCACAGCTCAGCTCAGGCGCT 7230

QY 201 AlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsn 220
 Db 7231 GCATTTCTGAACACCTGATGACCTCTCATCGCGGACCCCTCTTTGCGCACAGAAC 7290

QY 221 SerGlyLeuGlyGlnAsnGluThrLeuSerGlyArgAspIleGluHisHisLeuAlaAsn 240
 Db 7291 AGTGGGCTGGGAGATGAGACTTCTTGGGCGGACATCGAGCACCACTCGCAAT 7350

QY 241 ValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGly 260
 Db 7351 GTCAGCATGTTTCTTACAAATGACTTGTCAATGGCACCACTGCAACGAGGCTGGGA 7410

QY 261 SerLysLeuLeuIleThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal 280
 Db 7411 AGCAAGTGTCTCATCTGCGGAGAGAGCCCTTCCACCGAGGAGACAGGTTGTT 7470

QY 281 AspGlyArgAspThrLeuGluTyrAspIleCysAlaSerAsnGlyIleThrHisValIle 300
 Db 7471 GATGGAAGAGCCATTCTGCGAGTGGGACATCTTTGCCCTCAATGGATCATTCATGTCAT 7530

QY 301 SerArg***LeuLysAlaProProAlaProValThrLeu***HisThrGlyLeuGly*** 320
 Db 7531 TCCAGGCGCTTTAAAGCACCCCTTCCCGCGGAGTACCTTGACCCACACTGGCTGGGAGCA 7590

QY 321 GlyIlePhe*****IleLeuValThrGlyValAlaValAlaLeuAlaTyrSerTyr 340
 Db 7591 GGGATCTTCTTGCCATCACTGCTGAGTGGGCTGTGCTTGGCTGCTTACTCTCTAC 7650

QY 341 PheArgIleAsnArgLysThrIleGlyPhe***HisPhe 353
 Db 7651 TTTCGGATAAACCGAGAACATCGGCTTCCAGCATTTT 7689

RESULT 6
 HSA295695
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REMARK
 FEATURES

HSA295695 8266 bp mRNA linear PRI 28-JAN-2003
 Homo sapiens mRNA for stabilin-2 (STAB2 gene).
 AJ295695
 AJ295695.1 GI:18073427
 STAB2 gene; stabilin-2.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1
 Politz,O., Gratchev,A., McCourt,P.A.G., Schledzewski,K., Guillot,P., Johansson,S., Svineng,G., Franke,P., Kamnitsch,C., Kzyshkowska,J., Longati,P., Veltin,F.W., Johansson,S. and Goerdts,S.
 Stabilin-1 and -2 constitute a novel family of fasciclin-like hyaluronan receptor homologues
 Biochem. J. 362 (Pt 1), 155-164 (2002)
 2
 21818465
 11829752
 3
 Politz,O., Gratchev,A., McCourt,P.A.G., Schledzewski,K., Guillot,P., Johansson,S., Birk,R., Hakiy,N., Franke,P., Kodelja,V., Kamnitsch,C., Orfanos,C.E., Johansson,S. and Goerdts,S.
 Stabilin-1 and stabilin-2 constitute a novel family of fasciclin domain-containing adhesion molecules associated with endothelial-macrophage differentiation and angiogenic processes
 Unpublished
 3 (bases 1 to 8266)
 Gratchev,A.
 Direct Submission
 Submitted (27-JUL-2000) Gratchev A., Department of Dermatology, University Medical Center Mannheim, Theodor-Kutzer-Ufer 1, D-68135 Mannheim, GERMANY
 Revised by author 27-APR-2001
 Location/Qualifiers
 1..8266
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 1..8266
 /gene="STAB2"
 187..7842
 /gene="STAB2"
 /codon_start=1
 /product="stabilin-2"
 /protein_id="CAC82105.1"
 /db_xref="GI:18073428"
 /db_xref="GOA:Q8WQ8"
 /db_xref="SPTREMBL:Q8WQ8"

gene
 CDS

SCALNGLVKCPDGYTMTITSGSVGRDCTVFEVRYTSLSLGCRHICRKYLOPRCCP
 GWGPCICPGAGSPGNGRSCARGMEGNGTCSCQGGGTGACACADNLLFGPSC
 SSVCNCHVGVCSGLDGTCTCYSAITGPKDKPIPECAALLCPENSRCSTEDEN
 KLECKLPNTRGDGKYCDIPNCLRKICHPHACTYLGPNRHSCTCEGTRGGQVCL
 PVPDQINFGNCTKSTCKYDGGQSHCEKHEYNFVFGVSGWMDICKSNPCHR
 NANCITVAPRTCEICQKGVGDLTCYGNIMERLELNTEPRGKQGLTSLISLD
 KATFVTLTKGSGEIPNSDKDNOIKLHGKGVKVIIOGDIASNGLLHIDRAMDKL
 TQWYFLLTKGSGEIPNSDKDNOIKLHGKGVKVIIOGDIASNGLLHIDRAMDKL
 EHTFESNNGTMTMLQPRYSKPSRSLLETLNGLHALDEGVGPGPYTIVPNNALNM
 KUTGLYLLSPESRKLILVRYHVPTTQLEVALTILSTPHSRMANQLIQNTTNG
 QILANDVAMEEIEITAKNGRIYTLTLIPPSIVPLPHRDETKREMLKIGTCVCSL
 VYWSRFPANSEPTALFTRCVYSGRGLSKGCARYCNATVKIPCKGFGYGDQCQC
 PGFSNPGSCNGQADSLGNGTICIEGFGSQSCQFCDPNPKYPRCKKXKLCVHTGT
 CNNRIDCGACLTCTCRDGSAGRLCDKQTSACQPVQFCHIHATCEYNGTACICKA
 GYEGDCTLCSEMDPCTGLTPGSCSRNAECIKTGTGTHTCVCOQWNGNRDCSEIINC
 LLPASAGCHDNASCLYVPGQNECECKKFRGNGIDCEPITSCLEOTGKCHPLASQCS
 VLVPSQWSCVCGEGDGLCYGNAAVEISFLSEAAIFNRWINNALSQPTLSATSNLT
 VLVPSQWATEDMDQDEKFLWSQSNIPALIKYHMLLTGYRADLTSSLPLMLRLEQ
 GNPLHLAKVGNITIEGASIVDGNAAATNGVTHINKVLVPPQRLTGLSPNLMRLAQ
 MPDYPFRGYIIOYNLANALEADAVTFAPNNNAIENYIREKKVLSLEEDLVRYHV
 LEEKLAKNDLHNGMRETMGLGFSYFLSFFLHNDQLYNEAPINYTNVATDKGVHGLG
 KVLIEQKNRCDNNATTILIKRCKTCSSETCTCPFGTKSLGNEKRCRIYTSYFMGRRLIF

IGCPKCVRTVITRECCAFPGPQCPQCPGNAQNVCFNGICLDGVNVTGCEGEGF
SGTACETCTBGIYGHCDQACSVHRCNQGPDGSDCDVWGRVHCDNATIEDNC
NRVCHTASCLNTSDGTASCKAQFQNGTCTAINACEISNCGSCAKADCKRTTG
RRVCTCAGYTGDIIVCLIEINPCLENHGGCDKNAECTGTGPNACNCLPAYTDGKV
CTLINCLTNGKGGSEFAICNHTGQVERTCTCKPNIIGDGTFCRGSIVQELPKNPKS
QYFFQLQSHFVKGLVGPFPFTVFAPLSAARDEEARVKDWDKYGMPQVLRVHVACHQ
LLENLKLISNATSLQGFPIVSVQSTVYINNAKILSSDIISNGIHLIDKLLSP
KNLLITPKDNGRILQNTLATTNNGYIKFNSLIDQGLLSVIDPHTPVTILFWPTD
QALHALPAEQDFLFPQNDOKLKEYLKFHVIDAKVLAVDLPSTSTAWKTLQGSLSV
KCGARDIGDLFLNGQTCRIVQRELLFDLGVAYGIDCLLIDFTLGGRCDTFTTFDAG
ECGSCVNTSPSPRMSKPGVKOKCLYNLPKRNLEGGRCRSLVIOIPRCCKGVFGRD
COACPGGPDAPNNRGVCLDQYSATGECKQNTGNGTACEMCPGRFGPCLPCGSD
HQCDDITGSGQCLCETGTPGSCDQAVLPVACTPCSHATCKENNTCECNIDYE
GGCCTIVDFKQDNGCAVKARQKGTQVSCGQYKGGHSCREIDPCADGNL
GGCHETATGKMTGPKHCKESHVYDGLNCEPEQPIIDRLQDQCGHADAKVDL
HFQDTTGVFHLRSPLOQYKLTDFDKAREACENAAATWATYNQLSAQAKYHLGSAW
LETGRVAYTAFASQNGSGVGIIVDYGRPNKSEMMDVFCYRMKDVNCTCKGVVNGD
GFSGNTLQVLMFSPSLTNFLTEVLAYSNSARGRAFLHLTDLISRTGLTFVQNSG
LGENETLSGRDI EHLANVSMEFYNDLVNGTTLQTRLGSKLLITASQDPLQPTETRFV
DGRALLOWDIFASNGI IHVSRPLKAPAPVTLFHTGLGAGIFAILIVTGAVALAAY
SYFRINRTTIGFQHFSEEDINVAALGKQQPENISNPLYESTTSAPPEPSYDPTDSE
BRQLEGNDRPLTL"

ORIGIN

Alignment Scores:
Pred. No.: 5,65e-164 Length: 8266
Score: 1772.00 Matches: 334
Percent Similarity: 95.18% Conservative: 2
Best Local Similarity: 94.62% Mismatches: 17
Query Match: 95.01% Indels: 0
DB: 9 Gaps: 0

US-09-466-778B-11 (1-353) x HSA295695 (1-8266)

Qy 1 MetThrGlyProGlyLeuHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeu 20
Db 6613 ATGACAGCGCCGGGCAAGCAAGTGTGAGTGTAAAGTCACTATGTGGAGATGGGCTG 6672
Qy 21 AsnCysGluProGluGlnLeuProIleAspArgCysLeuGlnAspAsnGlnCysHis 40
Db 6673 AACTGTGAGCGGAGCAGCTGCCATTCACCGCTGCTTACAGGACAATGGGCACTGCCAT 6732
Qy 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60
Db 6733 CGAGACGCAAAATGTGTGACCTCCACTCCAGGTATACACTGTGGGGTGTTCACATCA 6792
Qy 61 ArgSerProLeuGlyGlnTyrIleLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80
Db 6793 CGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGCGCTGTGCCAAC 6852
Qy 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis 100
Db 6853 GAAGCTGGACCATGGCAACCTACACACGCTCTCTATGCCAGAGCCCAAGTACCAC 6912
Qy 101 LeuCysSerAlaGlyTyrIlePleuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120
Db 6913 CTGTGCTCAGCAGGCTGGCTGGAGACCGGGGGTGTGCTACCCACAGACCTTCGCCCTCC 6972
Qy 121 GlnAsnCysGlySerGlyValValGlylleValAspTyrGlyProArgProAsnLysSer 140
Db 6973 CAGAACTGTGGCTCTGTGTGGTGGATAGTGGACTATGGACCTAGACCTAGACCCCAAGAGT 7032
Qy 141 GluMetTyrAspValPheCysTyrArgMetLysAspValAsnCysThr***LysValGly 160
Db 7033 GAAATGTGGAGTGTCTTCGCTATCGGATGAAGATGTGAATGCACCTGCAAGTGGGC 7092
Qy 161 TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuLeuGlnValLeuMetSerPhePro 180
Db 7093 TATGTGGGAGATGGTCTTCATGCGAGTGGGAACCTGTCTGCAGGTCTCATGTCTTCCTCC 7152
Qy 181 SerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArg 200
Db 7153 TCACTCACAAACTTCTGACGGAAGTGTGGCTATTCCACAGCTCAGCTCAGCGCGGT 7212

Qy 201 AlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGluAsn 220
Db 7213 GCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGCACCTCTTTGTGCCACAGAAC 7272
Qy 221 SerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisLeuAlaAsn 240
Db 7273 AGTGGGCTGGGGGAGATGAGACCTGTCTGGGGGGGACATCGAGCACCACCTCGCCAAT 7332
Qy 241 ValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGly 260
Db 7333 GTACAGCATGTTTCTACATGACCTTGTCAATGCGACCCCTGCAACGAGGCTGGGA 7392
Qy 261 SerLysLeuLeuIleThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal 280
Db 7393 AGCAAGCTGTCTCATCACTGCCAGCAGCAGCACCACCTCCAAACCGAGGAGACCGAGTTTGT 7452
Qy 281 AspGlyArgAspThrLeuGluTyrAspIleCysAlaSerAsnGlyIleThrHisValIle 300
Db 7453 GATGGAAGACCATCTTCAGTGGGACATCTTTCCTCCATGAGATCATTCATGTCAAT 7512
Qy 301 SerArg***LeuLysAlaProProAlaProValThrLeu***HisThrGlyLeuGly*** 320
Db 7513 TCCAGGCTTTTAAAGCACCCCTGCCCCGTGACCTTGACCCACACTGGCTTGGGAGCA 7572
Qy 321 GlyIlePhe***IleLeuValThrGlyAlaValAlaLeuAlaAlaTyrSerTyr 340
Db 7573 GGGATCTTCTTGTGCATCATCTGTGACTGGGGCTGTGCTTGGTGTCTTACTCTCTAC 7632
Qy 341 PheArgIleAsnArgLysThrIleGlyPhe***HisPhe 353
Db 7633 TTTCCGATTAACCGAGACAATCGGCTTCCAGCATTTT 7671

RESULT 7

AX686401 LOCUS 3260 bp DNA linear PAT 29-MAR-2003
DEFINITION Sequence 210 from Patent WO02059315.
ACCESSION AX686401
VERSION AX686401.1 GI:29372140
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Shimkets,R.A., Patturajan,M., Vernet,C.A., Casman,S.J.,
Malyanekar,U., Shenoy,S., Spytek,K.A., Gangolli,E., Miller,C.,
Baldog,F., Li,L., Taupier,R.J., Kekuda,R., Smithson,G.,
Zerhusen,B.D., Liu,X., Colman,S.D., Tchernev,V., Si,J., Edinger,S.,
Stone,D., Sciore,P., Millet,I. and Rothenberg,M.
Human nucleic acids and polypeptides and methods of use thereof
Patent: WO 02059315-A 210 01-AUG-2002;
Curagen Corporation (US)
FEATURES
Location/Qualifiers
source
1..3260
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 3,44e-164 Length: 3260
Score: 1769.00 Matches: 333
Percent Similarity: 95.18% Conservative: 3
Best Local Similarity: 94.33% Mismatches: 17
Query Match: 94.85% Indels: 0
DB: 6 Gaps: 0

US-09-466-778B-11 (1-353) x AX686401 (1-3260)

Qy 1 MetThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeu 20
Db 1635 ATGACAGCGCCGGGCAAGCAAGTGTGAGTGTAAAGTCACTATGTGGAGATGGGCTG 1694

```

QY 21 AsnCysGluProGluGlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis 40
DB 1695 AACTGTGAGCGGAGCAGTGCCTTACCGCTGCTTACAGGACAATGGCAGTGCAT 1754

QY 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60
DB 1755 GCAGACGCCAAATGTGCGACCTCCACTCCAGATACCACTGTGGGTGTTCATCTA 1814

QY 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80
DB 1815 CGCTCCCACTGGCCAGATATAGCTTACCTTACAAAGCCAGAGGCTGTGCCAAC 1874

QY 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis 100
DB 1875 GAAGCTCGCACCATGGCAACCTCAACACAGCTCTCTATGCCAGAAAGGCCAAGTACAC 1934

QY 101 LeuCysSerAlaGlyTyrLeuGluThrGlyArgValAlaLysProThrAlaPheAlaSer 120
DB 1935 CTGTGCTCAGCAGGCTGGCTGGAGACCGGGGGGTTCCTACCCCAAGCTTCGCCTCC 1994

QY 121 GlnAsnCysGlySerGlyValValGlyLysValAspTyrGlyProArgProAsnLysSer 140
DB 1995 CAGAACTGTGGCTCTGTGTGGTGGTGGTAGTGGACTATGACCTAGACCAACAAGAGT 2054

QY 141 GluMetTrpAspValPheCysTyrArgMetLysAspValAsnCysThr***LysValGly 160
DB 2055 GAAATGTGGGATGTCTTCTGCTATCGGATGAAGATGTGAACCTGCACCTGCAGGTGGC 2114

QY 161 TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuGlnValLeuMetSerPhePro 180
DB 2115 TAGTGGGAGATGGCTTCTCATGAGTGGGAACTGCTGAGGTCTTGATGCTCTTCCCC 2174

QY 181 SerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArg 200
DB 2175 TCATCTCAAACTTCTCTGACGGAAGTGTGCTTATTCACACAGCTCAGCTGCAGGTGGC 2234

QY 201 AlaPheLeuGlnHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsn 220
DB 2235 GCATTTCTAGAACACCTGACTGACTCTGCTTCCATCGCGGACCTCTTTGTGCCACAGA 2294

QY 221 SerGlyLeuGlyGlnAsnGluThrLeuSerGlyArgAspIleGluHisLeuAlaAsn 240
DB 2295 AGTGGGCTGGGGAGATAGACCTTGTCTGGGCGGACATCGAGCACCACTCGCCAAAT 2354

QY 241 ValSerMetPhePheThrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGly 260
DB 2355 GTCAGCATGTTTCTTACATGACTTGTCAATGGCACCCCTGCAACAGAGGTGGGA 2414

QY 261 SerLysLeuLeuIleThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal 280
DB 2415 AGCAAGCTGCTCATCATCTGCAGCGACGAGCCCACTCCCAACCGAGGAGACGAGTTTGT 2474

QY 281 AspGlyArgAspThrLeuGluTrpAspIleCysAlaSerAsnGlyIleThrHisValIle 300
DB 2475 GATGGAGAGACCATCTGCGAGTGGGACATCTTTCCTCCATGGATGCATTCATGTCAT 2534

QY 301 SerArg***LeuLysAlaProProAlaProValThrLeu***HisThrGlyLeuGly*** 320
DB 2535 TCCAGGCTTTAAAGCACCCCTGCTCCCGTGCCTTGCACCTGACCCACACTGGCTGGGAGCA 2594

QY 321 GlyIlePhe*****IleLeuValThrGlyAlaValAlaLeuAlaLysSerTyr 340
DB 2595 GGGATCTCTTGTGCATCATCTGCTGACTGGGCTGTTCCTTGGCTGCTTACTCTCTAC 2654

QY 341 PheArgIleAsnArgLysThrIleGlyPhe***HisPhe 353
DB 2655 TTTCGGATAAACCGAGAACATCGCTTCCAGCATTTT 2693

```

RESULT 8

AF160476

LOCUS

DEFINITION Homo sapiens CD44-like precursor FELL mRNA, linear PRI 12-JUL-2000

ACCESSION

AF160476.1 GI:9049505

VERSION

Homo sapiens (human)

KEYWORDS

Homo sapiens

SOURCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 3260)

AUTHORS

Tao, Q., Zhang, W. and Cao, X.

TITLE

Molecular cloning and characterization of human FELL sharing

JOURNAL

Unpublished with CD44

REFERENCE

2 (bases 1 to 3260)

AUTHORS

Tao, Q., Zhang, W. and Cao, X.

TITLE

Direct Submission

JOURNAL

Submitted (21-JUN-1999) Department of Immunology, Shanghai

Brilliance Biotechnology Institute & Second Military Medical

University, 800 Xiangyin Road, Shanghai 200433, P. R. China

FEATURES

Location/Qualifiers

1..3260

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

171..2864

/codon_start=1

/product="CD44-like precursor FELL"

/protein_id="AAF82398.1"

/db_xref="GI:9049506"

/translation="MPQVLRVHVACHQLLENKLSNATSLQGPPIVSVSQSTVY

INNKAKISDIIISNGIHLIDKLLSPKNLLITPKDNRSLILQTLTATNNGYIKF

SNLIQSGLLSVITDTHIPVTLFWPTDQALHALPAEQDFLENQDNKKLYLKFH

VIRDAKLVADLPTSTAWKTLQGSLSVKGAGRDIDGDFLNGQTRIVORELLFDIG

VAYIDCLLDITLGGRCITFTTFFDASGECSCVNTSPCRWSKPKGVKKLYNLFP

KRLEGGRCRCSLVIQIPRCKGYFGRCQACPGPDAPCNRRGVCLIDQVTSATGECK

NTGFNCTACEMWPFGRDCLPCGSCDHGQDDGTGSGQCLCETGWTGPGSDTOAV

LPVCTTPGCSAHATKENNTCECNLDYEGDITCTVDFCKDNGSCAKVARCSQGT

KSCSCQKGYKGCCHSCTEIDPCADGLNGGCHEHATCKMTGPKGKCECKSHVYDGL

NCSEPOLIDRCLQDNGQCHADKVDLHFQDTTGVFHLRSLPGYKLTFFDKAREAC

ANEAATMYATNQLSYAKAKYHLCSAGWLETGRVAYPTAFASQNGSGVGVIVDYP

PNKSEMDVFCVRMKDVTCKVYVGDGFCGSLNLLQVMSPPSLNFTLFTVLYSNG

SSARGAFLEHLTDLISIRGTLFVPSNGLGENETLSGRDIEHLANVSMPFYNDLVNG

TLTOTVSGKLLITASODPLQPTETFPVGRALIWQDI PASNGI IHIISRPKAPAP

VTLTHGLAGLGIFFAILVLTGVALAAAYSYFRINRTTIGTFQHFESIEDINVAALGKQ

PENISNPLYESITTSAPPEPSYDPTDSEERQLEGNDPLRLT"

ORIGIN

Alignment Scores:

Pred. No.: 3 44e-164 Length: 3260
 Score: 1769.00 Matches: 333
 Percent Similarity: 95.18% Conservative: 3
 Best Local Similarity: 94.33% Mismatches: 17
 Query Match: 94.85% Indels: 0
 DB: 9 Gaps: 0

US-09-466-778B-11 (1-353) x AF160476 (1-3260)

QY 1 MetThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeu 20
 DB 1635 ATGACAGGCGCGGCAAGCACAAAGTGTGAGTGTAAAGTCACTATGTCGAGATGGGCTG 1694
 QY 21 AsnCysGluProGluGlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis 40
 DB 1695 AACTGTGAGCGGAGAGCTGCCCATTTGACCGCTGCTTACAGGACAATGGCAGTGCAT 1754
 QY 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60
 DB 1755 GCAGACGCCAAATGTGCGACCTCCACTTCCAGATACCACTGTGGGTGTTCATCTA 1814
 QY 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80
 DB 1815 CGCTCCCACTGGCCAGTATAGCTTGTGACCTTTGACAAAGCCAGAGGCTGTGCCAAC 1874
 QY 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis 100

```
Db 1875 GAAGCTGGACCATGGCAACCTACCAACAGCTCTCTATGCCCCGAAGCCCAAGTACCAC 1934
Qy 101 LeuCySerAlaGlyTyrLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120
Db 1935 CTGTGCTCAGCAGGCTGGCTGGAGCCGGGGTGGCTTACCCACAGCCCTTGCCTCC 1994
Qy 121 GlnAsnCysGlySerGlyValValGlyIleValAspTyrGlyProArgProAsnLysSer 140
Db 1995 CAGAACTGGCTCTGGTGTGGTGGAGTAGTGGACTATGACCTAGACCCCAACAAGAGT 2054
Qy 141 GluMetTyrAspValPheCysTyrArgMetLysAspValAsnCysThr***LysValGly 160
Db 2055 GAAATGTGGATGCTCTCTGCTATCGGATGAAGATGGAATGCACCTGCAAGGTGGC 2114
Qy 161 TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuLeuGlnValLeuMetSerPhePro 180
Db 2115 TATGTGGAGATGGCTTCTCATGCASTGGAACTGTCTGCAGGTCTGATGTCTTCCCC 2174
Qy 181 SerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArg 200
Db 2175 TCACCTACAAAATTCTCTCAGCGAAGTGTGGCTATTCCAAACAGCTCAGCTCGAGCCGT 2234
Qy 201 AlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsn 220
Db 2235 GCATTTCTAGAACACCTGACTGACCTGTCCATCCCGCGGACCCCTCTTTGTGCCACAGAAC 2294
Qy 221 SerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisLeuAlaAsn 240
Db 2295 AGTGGCTGGGGAGAAATGAGACCTGTCTGGCGGGACATCGAGCACCACTCGCCAAAT 2354
Qy 241 ValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGly 260
Db 2355 GTCAGCATGTTTCTACAAATGACCTGTGCAATGGCACCACTCCAAACAGAGGGGTGGGA 2414
Qy 261 SerLysLeuLeuIleThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal 280
Db 2415 AGCAAGCTGCTCATCCTGCGAGCCAGGACCCACCTCCCAACGAGGAGACCAAGGTTGTT 2474
Qy 281 AspGlyArgAspThrLeuGluTrpAspIleCysAlaSerAsnGlyIleThrHisValIle 300
Db 2475 GATGGAAGAGCATCTCGAGTGGGACATCTTTGCTTCCAATGGGATCATTCATGTCATT 2534
Qy 301 SerArg***LeuLysAlaProProAlaProValThrLeu***HisThrGlyLeuGly*** 320
Db 2535 TCCAGGCTTTTAAAGACACCCCTCCCTCCCGCTGACCTTTGACCCACACTGGCTTGGAGCA 2594
Qy 321 GlyIlePhe*****IleLeuValThrGlyAlaValAlaLeuAlaAlaTyrSerTyr 340
Db 2595 GGGATCTTCTTTGCCATCATCTGCTGACTGGGGCTGTGCTTGGCTTGTCTTACTTCCAC 2654
Qy 341 PheArgIleAsnArgLysThrIleGlyPhe***HisPhe 353
Db 2655 TTTCCGATAAACCGAGAACAAATCGCTTCCAGCATTTT 2693

RESULT 9
AX149461
LOCUS AX149461 3625 bp DNA linear PAT 08-JUN-2001
DEFINITION Sequence 17 from Patent WO0136638.
ACCESSION AX149461
VERSION AX149461.1 GI:14347936
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
Shimkets, R.A., Lichenstein, H., Vernet, C. and Fernandes, E.
AUTHORS Polypeptides and nucleic acids encoding same
TITLE Patent: WO 0136638-A 17 25-MAY-2001;
JOURNAL Curagen Corporation (US)
FEATURES
Location/Qualifiers
1..3625
source
```

```
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores: 2.48e-143 Length: 3625
Pred. NO.: 1558.00 Matches: 300
Score: 85.31% Conservative: 2
Percent Similarity: 84.75% Mismatches: 15
Best Local Similarity: 83.54% Indels: 37
Query Match: 6 Gaps: 1
DB: 1

US-09-466-778b-11 (1-353) x AX149461 (1-3625)

Qy 1 MetThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeu 20
Db 2078 ATGACAGGCCCGGGCAAGCACAGTGTGAGTGTAAGAGTCACATATGTCCGGAGATGGGCTG 2137
Qy 21 AsnCysGluProGluGlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis 40
Db 2138 AACGTGTAGCGCGGAGCAGCTGCCCATTTGACCGCTGCTTACAGGACAAATGGGACGTGCAT 2197
Qy 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisIleu 60
Db 2198 GCAGACGCCAAATGTGTGACCTCCACTTCAGGATACCACTGTGTGGGTGTTCCTATCTA 2257
Qy 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80
Db 2258 CGCTCCCACTGGGCCAGTATAGCTGACCTTTCACAAAGCCAGAGGCTGTGCCAAC 2317
Qy 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis 100
Db 2318 GAAGCTCGACCATGGCAACCTACACAGCTCTCTATGCCAGAGGCGCAAGTACCAC 2377
Qy 101 LeuCySerAlaGlyTyrLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120
Db 2378 CTGTGCTCAGCAGGCTGGCTGGAGCCGGGGTGGCTTACCCACAGCCCTTGCCTCC 2437
Qy 121 GlnAsnCysGlySerGlyValValGlyIleValAspTyrGlyProArgProAsnLysSer 140
Db 2438 CAGAACTGTGGCTCTGGTGTGGTGGATAGTGGACTATGACCTAGACCCCAACAAGAGT 2497
Qy 141 GluMetTyrAspValPheCysTyrArgMet-LysAspValAsnCysThr***LysValGly 160
Db 2498 GAAATGTGGATGTCTTCTGCTATCGATGAAAG----- 2531
Qy 160 yTyrValGlyAspGlyPheSerTyrSerGlyAsnLeuLeuGlnValLeuMetSerPhePr 180
Db 2531 ----- 2531
Qy 180 oSerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyAr 200
Db 2532 -----GAAGTGTGGCTATTTCACACAGCTCAGCTCGAGGCG 2569
Qy 200 gAlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAs 220
Db 2570 TGCATTTCTAGAACACCTGACTGACCTGCCATCCCGCGGACCCCTCTTTGTGCCACAGAA 2629
Qy 220 nSerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisHisLeuAlaAs 240
Db 2630 CAGTGGGCTGGGGAGAAATGAGACCTTGTCTGGCGGGGACATCGAGCACCACTTCGCCAA 2699
Qy 240 nValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGl 260
Db 2690 TGTGAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCCACCTCGCAACAGGCTGGG 2749
Qy 260 ySerLysLeuLeuIleThrAspArgGlnAspProLeuHisProThrGluThrArgCysVa 280
Db 2750 AAGCAAGCTGCTCATCTACCTCCAGCCAGGACCCCACTCCAAACCGAGGAGACCAAGTTGT 2809
Qy 280 lAspGlyArgAspThrLeuGluTrpAspIleCysAlaSerAsnGlyIleThrHisValIl 300
Db -----
```



```
Db 2810 TGATGGAAGACCAATTCGAGTGGGACATCTTTGGCTCCAAATGGGATCATCTCATGTGCAT 2869
Qy 300 eSerArg***LeuylsAlaProProAlaProValThrLeu***HisThrGlyLeuGly** 320
Db 2870 TTCCAGGCGCTTTAAAGACACCCCTGCGCCCGTGCACCTTGACCCACACTGGCTGGGAGC 2929
Qy 320 *GlyLeuPhe*****llelleuValThrGlyAlaValAlaLeuAlaValThrSerTy 340
Db 2930 AGGAGATCTTTTGGCATATCTCCGATGCTGGGCTGTTGGCTGGCTGCTTACTCCTA 2989
Qy 340 rPheArgIleAenArgIleThrIleGlyPhe***HisPhe 353
Db 2990 CTTTCGGATTAACCGGAGACATATCGGCTCCAGCATTTT 3029

RESULT 10
HSM801377
LOCUS
DEFINITION Homo sapiens mRNA; cDNA DKFZp434E0321 (from clone DKFZp434E0321).
ACCESSION AL133021
VERSION AL133021.1 GI:6453577
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS Blum, H., Bauersachs, S., Mewes, H. W., Gassenhuber, J. and Wiemann, S.
TITLE Direct Submission
JOURNAL Submitted (15-NOV-1999) MIPS, Am Klopferspitze 18a, D-82152
Martinsried, GERMANY
COMMENT
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
sequenced by IMU (Ludwig Maximilians University,
Munich/Germany) within the cDNA sequencing consortium of the German
Genome Project.
This clone (DKFZp434E0321) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://www.mips.biochem.mpg.de/proj/cDNA/.
FEATURES
source
1..3642
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp434E0321"
/tissue_type="testis"
/clone_lib="434 (synonym: htes3). Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
join(7)..2607,2609..3217
/gene="DKFZp434E0321"
join(<7)..2607,2609..3217
/gene="DKFZp434E0321"
/notes="strong similarity to KIAA0246
The frame shift was determined manually"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAB61358.1"
/db_xref="GI:6453578"
/db_xref="GOA:Q9UF98"
/db_xref="SPTRMBL:Q9UF98"
/translation="VGEAVGTASCKCAAGFGNGTICTAINACEISNGGCSAKADCKR
TTPRRVCTKAGTVDGIVGICLENPCLEHGGCDKNAECTQTGPNOAACNCLPVTG
DGKVTILNVLTKNGCSFAICNHTGQVETCTCKPNYIGDGTGRGSIYDELPHV
PKTSQYFQLEHFVLDVGPFTVAPLSAADFBEARVKDWKGLMPQVLRVHV
ACHOLLENKILSNATSLQGEPIVISVOSTVYINNKAIISSDIISTWGIYHIDK
LLSPKNLIITPKNSGRILNLTATNNGYIKFNSNLQDSGLLSVITDPIHTPVILF
WPTDQALHAPASQDDFLNQDNKDKLKEYLKHVIRDAKVLAVDLPSTFAWTKLOS
ELSVKCGAGRIDGLFLNGQTCIRVQRELLFDLGVAYGIDCLIDPLTGRCDTFTF
DASGECGSVNTSCPSRWSPKGVQKCLNLPKRLNLEGRERCSLVIQIPRCCKGY
FGRDQACPGPAPCPNRRGVCLDOYSATGCKNTGFCGACEMCPWGRFGRDCLPC
GCSDHGQDDGIGTSGGQCLCTGWTGFSCTQAVLPVCTPPCSAHATCKENNTCEGN
```

```
LDYEGDITCTVVDFCKQNGGCAKVARCSQKTKVSCSQKQKYGKGDHSCITDPCA
DGLNGCHEHATCKMTGPKHCKECSHYVGLNCEPEQLPIRDLQNGQCHADAK
CYDLHFQDITTVGVFHLRSPLGQYKLTFDKAREACANEATMATYNQLSYAKAKYHLC
SAGLWETGVAVPTAFASQNGSGVVGIVDYGRPNKSEMDVFCYRMKSGAGLFOOL
SRPCTSRTPDDLSIRGTLFVPONSGLGENETLSGROIHHLIANSMFVYNDLVNMT
LQTRLSKLLITASODPLOTPTFRFVDSGAILQWDIPASNGLIHVSRLKAPPAVPT
LHTGUGAGIFAILIVTGAVALAAYSIFRINRTTIGFQHFSEEDINVAALQKQOPE
NISNPLEYESTTSAPPEPSYDPFTDSEERQLEGNDRPLRLT"
polyA signal 3574..3579
polyA_site 3603
ORIGIN
Alignment Scores:
Pred. No.: 2,49e-143 Length: 3642
Score: 1558.00 Matches: 300
Percent Similarity: 85.31% Conservative: 2
Best Local Similarity: 84.75% Mismatches: 15
Query Match: 83.54% Indels: 37
DB: 9 Gaps: 1
US-09-466-778B-11 (1-353) x HSM801377 (1-3642)
Qy 1 MetThrGlyProGlyVlyeHisCysGluCysAlaSerHisTyrValGlyAspGlyLeu 20
Db 2095 ATGACAGGCCCGGCGAAGCAAGTGTGAGTGTAAAGTCACATGTCGAGATGGCTG 2154
Qy 21 AsnCysGluProGluInLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis 40
Db 2155 AACTGTGAGCGGAGCAGCTGCCCAITGACCGCTGTACAGCAATGGCAGTGCAT 2214
Qy 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60
Db 2215 GCAGAGCCCAATGTGTGCACCTCCACTTCCAGATACCACTGTGGGGTGTTCATCTA 2274
Qy 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80
Db 2275 CGTCCCGACCTGGCGCAGTATAAGCTTGACCTTTGACAAAGCCAGAGAGGCTGTGCCAAC 2334
Qy 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis 100
Db 2335 GAAGCTGCGACCATGCAACCTACACAGCTCTCTCTATGCCAGAGGCGCAAGTACCAC 2394
Qy 101 LeuCysSerAlaGlyTrpLeuGluThrGlyArgValAlaLysThrProThrAlaPheAlaSer 120
Db 2395 CTGTGCTCAGCAGCGTGGCTGGAGACGGGCGGGTGGCTTACCCACAGCTTTCGCTCC 2454
Qy 121 GlnAsnCysGlySerGlyValValGlyValValAspTyrGlyProArgProAsnLysSer 140
Db 2455 CAGAACTGTGGCTCTGGTGTGGTGGATAGTGGACTATGGACTAGACCCACACAGAGT 2514
Qy 141 GluMetTrpAspValPheCysTyrArgMet-LysAspValAsnCysThr***LysValGln 160
Db 2515 GAAATGTGGCATGCTCTCTGCTATCGATGAAAG----- 2548
Qy 160 yTyrValGlyAspGlyPheSerTyrSerGlyAsnLeuLeuValLeuMetSerPhePr 180
Db 2548 ----- 2548
Qy 180 oSerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyAr 200
Db 2549 -----GAAAGTGTGGCTATTTCACAGCTCAGCTCAGCGGCGG 2586
Qy 200 gAlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAs 220
Db 2587 TGCATTTCAGAACACCTGACTGACCTCCATCCCGGCGGACCCCTCTTTGTGCGCACAGAA 2646
Qy 220 nSerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisLeuAlaAs 240
Db 2647 CAGTGGGCTGGGGAGAAATGAGACCTTGTCTGGGCGGAGCATCGACGACACCATCGCCAA 2706
Qy 240 nValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGln 260
Db 2707 TGTACAGCATGTTTTCTTACAAATGACCTTGTCAATGGCACCCCTCTGAAACGAGGCTGGG 2766
```

```

Qy 260 ySerIysLeuLeuLeuThrAspArgGlnAspProLeuHisProThrGluThrArgCysVa 280
Db 2767 AAGCAAGCTGCTCATCTGCGAGCCAGGACCACTCCAAACCGAGGAGACCCAGGTTTGT 2826
Qy 280 lAspGlyArgAspThrLeuGluThrAspIleCysAlaSerAsnGlyIleThrHisVal11 300
Db 2827 TGATGGAGAGCCATTGCGAGTGGGACATCTTTGCCCTCCAAATGGGATCATTGTCAT 2886
Qy 300 eSerArg***LeuLysAlaProProAlaProValThrLeu***HisThrGlyLeuGly** 320
Db 2887 TTCCAGGCGCTTTAAAGACACCCCTGCCCGCGTACCTTGACCCACACACTGGCTTGGGAGC 2946
Qy 320 *GlyIlePhe*****IleIleLeuValThrGlyAlaValAlaLeuAlaIleThrSerTy 340
Db 2947 AGGGATCTTCTTTGTCATCATCTGCTGACCTGGGCTGTTGCCCTGGCTGCTTACTCTTA 3006
Qy 340 rPheArgIleAsnArgLysThrIleGlyPhe***HisPhe 353
Db 3007 CTTTCGGATAAACCGGAGAACAAATCGGCTTCAGCATTTT 3046

RESULT 11
AX686194 8495 bp DNA linear PAT 29-MAR-2003
LOCUS Sequence 3 from Patent WO02059315.
DEFINITION AX686194
ACCESSION AX686194
VERSION AX686194.1 GI:29372027
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS Shimkets,R.A., Patturajan,M., Vernet,C.A., Casman,S.J.,
Malyankar,U., Shenoy,S., Spytek,K.A., Gangolli,E., Miller,C.,
Boldog,F., Li,L., Taupier,R.J., Kekuda,R., Smithson,G.,
Zerhusen,B.D., Liu,X., Colman,S.D., Tchernev,V., Si,J., Edinger,S.,
Stone,D., Sciore,P., Millet,I. and Rothenberg,M.
TITLE Human nucleic acids and polypeptides and methods of use thereof
JOURNAL Patent: WO 02059315-A 3 01-AUG-2002;
Curagen Corporation (US)
FEATURES
source
1..8495
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 7,27e-143 Length: 8495
Score: 1558.00 Matches: 300
Percent Similarity: 85.31% Conservative: 2
Best Local Similarity: 84.75% Mismatches: 15
Query Match: 83.54% Indels: 37
DB: 6 Gaps: 1

US-09-466-778B-11 (1-353) x AX686194 (1-8495)

Qy 1 MetThrGlyProGlyLysHisLysCysGluCysLysSerHisTyValGlyAspGlyLeu 20
Db 6948 ATGACAGGCGCGGCAAGCAAGTGTGAGTGTAAAGTCACTATGTCGAGATGGGCTG 7007
Qy 21 AaCysGluProGluGluLeuProIleAspArgCysLeuGlnAsnGlyGlnCysHis 40
Db 7008 AACTGTGAGCGCGAGCAGCTGCCCATGACCTGCTTACAGGACAATGGGCGAGTGCCT 7067
Qy 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60
Db 7068 GCAGAGCGCAATGTGCGACCTCCACTCCAGGATACCACTGTGGGGTGTCCACTTA 7127
Qy 61 ArgSerProLeuGlyGlnTyrlLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80
Db 7128 CGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACMAAGCCAGAGAGCGCTGTGCCAAC 7187

```

```

Qy 81 GluAlaIleThrMetAlaIleThrTyAsnGlnLeuSerTy***GlnLysAlaLysTyHis 100
Db 7188 GAAGCTCGGACCATGGCAACCTACCAACCACTCTCTATCCCAAGAGGCGCAAGTACCA 7247
Qy 101 LeuCysSerAlaGlyTyrLeuGluThrGlyArgValAlaIleThrProThrAlaPheAlaSer 120
Db 7248 CTGTGCTCAGCAGGCTGGCTGGAGACCGGGGGGTGCCTACCCCAAGGCTTCCCTCC 7307
Qy 121 GlnAsnCysGlySerGlyValValGlyIleValAspTyrlGlyProArgProAsnLysSer 140
Db 7308 CAGAACTGTGGCTCTGTGTGTGGATAGTGGACTATGGACCTAGACCAACCAAGAGT 7367
Qy 141 GluMetTrpAspValPheCysTyrlArgMet-LysAspValAsnCysThr***LysValGI 160
Db 7368 GAAATGGGATGTCTTCTGTCTATCGGATGAAAG----- 7401
Qy 160 yTyrValGlyAspGlyPheSerTyrlSerGlyAsnLeuLeuGlnValLeuMetSerPhePr 180
Db 7401 ----- 7401
Qy 180 oSerLeuThrAsnPheLeuThrGluValLeuAlaIleThrSerAsnSerSerAlaArgGlyAr 200
Db 7402 -----GAAAGTGTGGCTTATTCACACAGCTCAGCTCGAGGCGG 7439
Qy 200 gAlaPheLeuGluHisLeuThrAspLeuSerIleAtrGlyThrLeuPheValProGlnAs 220
Db 7440 TGCATTTCTAGAACACCTGACTGACCTGTCCATCCCGCGGACCCCTCTTTGTGCCACAGAA 7499
Qy 220 nSerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisLeuAlaAs 240
Db 7500 CAGTGGGCTGGGAGAGATGAGACCTTGTCTGGGCGGACATCGAGCACACCTTCGCCAA 7559
Qy 240 nValSerMetPhePheTyrlAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGI 260
Db 7560 TGTCCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACTCGCAACCGAGGCTGG 7619
Qy 260 ySerIysLeuLeuIleThrAspArgGlnAspProLeuHisProThrGluThrArgCysVa 280
Db 7620 AAGCAAGCTGCTCATCATCTGCGAGGAGGACCCCTCAACCGGAGGAGGACGAGTTTGT 7679
Qy 280 lAspGlyArgAspThrLeuGluThrAspIleCysAlaSerAsnGlyIleThrHisVal11 300
Db 7680 TGATGGAGAGCCATTCTGCGAGTGGACATCTTTGCTCCAAATGGGATCATTCATGTCAT 7739
Qy 300 eSerArg***LeuLysAlaProProAlaProValThrLeu***HisThrGlyLeuGly** 320
Db 7740 TTCCAGGCGCTTTAAAGACACCCCTGCCCGCTGACCTTGACCCACACTGGCTTGGGAGC 7799
Qy 320 *GlyIlePhe*****IleIleLeuValThrGlyAlaValAlaLeuAlaIleThrSerTy 340
Db 7800 AGGGATCTTCTTTCCTCCATCATCTGCGAGTGGGCTGTTGCCCTGGCTGCTTACTCTTA 7859
Qy 340 rPheArgIleAsnArgLysThrIleGlyPhe***HisPhe 353
Db 7860 CTTTCGGATAAACCGGAGAACAAATCGGCTTCAGCATTTT 7899

RESULT 12
AX686192 8444 bp DNA linear PAT 29-MAR-2003
LOCUS Sequence 1 from Patent WO02059315.
DEFINITION AX686192
ACCESSION AX686192
VERSION AX686192.1 GI:29372026
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 Shimkets,R.A., Patturajan,M., Vernet,C.A., Casman,S.J.,
Malyankar,U., Shenoy,S., Spytek,K.A., Gangolli,E., Miller,C.,
Boldog,F., Li,L., Taupier,R.J., Kekuda,R., Smithson,G.,
Zerhusen,B.D., Liu,X., Colman,S.D., Tchernev,V., Si,J., Edinger,S.,

```


Stone, D., Sciore, P., Millet, I. and Rothenberg, M.
Human nucleic acids and polypeptides and methods of use thereof
Patent: WO 02059315-A 1 01-AUG-2002;
Curagen Corporation (US)

FEATURES

Location/Qualifiers
1..8444
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 1 79e-142 Length: 8444
Score: 1554.00 Matches: 296
Percent Similarity: 85.27% Conservatives: 5
Best Local Similarity: 83.85% Mismatches: 16
Query Match: 83.32% Indels: 36
DB: 6 Gaps: 1

US-09-466-778B-11 (1-353) x AX586192 (1-8444)

Qy 1 MetThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeu 20
Db 6907 ATGACAGCGCCGGGCAACACAGTGTGAGTGTAAGTCACTATGTCGAGATGGGTG 6966

Qy 21 AsnCysGluProGluGlnLeuProLeuAspArgCysLeuGlnAsnGlyGlnCysHis 40
Db 6967 AACTGTGAGCGGAGCAGTCCCATTCACCGCTGCTTACAGGACAAATGGGCGAGTGCAT 7026

Qy 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60
Db 7027 GCACAGCGCAATGTGCGACTCCACTTCAGGATACCATGTGGGTGTTCATCTA 7086

Qy 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80
Db 7087 CGCTCCCACTGGGCGCACTAAGTGTACCTTTGACAAAGCCAGAGAGCGCTGTGCCAAC 7146

Qy 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis 100
Db 7147 GAAGCTGGGACCATGGCAACCTACACAGCTCTCTATGCCAGAGCAAGCCAGTACAC 7206

Qy 101 LeuCysSerAlaGlyTyrLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120
Db 7207 CTGTGCTCAGCAGGCTGGCTGGAGACCGGGGTGCTTACCCACAGCTTGCCTCC 7266

Qy 121 GlnAsnCysGlySerGlyValValGlyLeuValAspTyrGlyProArgProAsnLysSer 140
Db 7267 CAGAACTGTGGCTCTGGGTGGTGGATAGTGACTATGGACCTAGACCCACCAAGAGT 7326

Qy 141 GluMetTrpAspValPheCysTyrArgMetLysAspValAsnCysThr***LysValGly 160
Db 7327 GAAATGTGGATGTCTTCTGCTATCGGATGAAG----- 7359

Qy 161 TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuLeuGlnValLeuMetSerPhePro 180
Db 7359 ----- 7359

Qy 181 SerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArg 200
Db 7360 -----GAACTGTGGCTTATCCACAGCTATCCACAGCTCAGCTCGAGCGGT 7398

Qy 201 AlaPheLeuGluHisLeuThrAspLeuSerLeuArgGlyThrLeuPheValProGlnAsn 220
Db 7399 GCAATTTCTAGAACACCTGACTGACCTGCTCCATCCGCGGACCCCTCTTTGTGCGACAGAAC 7458

Qy 221 SerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAsnIleGluHisLeuAlaAsn 240
Db 7459 AGTGGCTGGGGGGAATGAGACCTTGTCTGGCGGGACATCGACACACCTCCCAAT 7518

Qy 241 ValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGly 260
Db 7519 GTCCAGCATGTTTCTACAACTGACCTGTTCATGGCACACCTCCGACAGGCTGGGA 7578

Qy 261 SerLysLeuLeuIleThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal 280
Db 7579 AGCAAGCTGTCTATCACTGCACGAGCCAGCCACTCCAAACGGGTACAAAGTAGTGTGT 7638

Qy 281 AspGlyArgAspThrLeuGluTyrAspIleCysAlaSerAsnGlyIleThrHisValIle 300
Db 7639 GATGGAAGAGCCATTCTGAGTGGGACATCTTGGCTCCATGGGATCATCATGTCA 7698

Qy 301 SerArg***LeuLysAlaProProAlaProValThrLeu***HisThrGlyLeuGly*** 320
Db 7699 TCCAGGCGCTTTAAAGACACCGCTGCGCCCTGACCTTGACCCACACCTGGCTGGGAGCA 7758

Qy 321 GlyIlePhe*****IleIleLeuValThrGlyAlaValAlaLeuAlaTyrSerTyr 340
Db 7759 GGGATCTCTTGTGATCATCTCTGGTACTGGGGCTGTGCTTGGCTTACTCTCTAC 7818

Qy 341 PheArgIleAsnArgLysThrIleGlyPhe***HisPhe 353
Db 7819 TTTGGGATAAACCGGAGAACAAATCGGCTACGACATTTT 7857

RESULT 13
AY007370 4706 bp mRNA linear ROD 23-OCT-2002
LOCUS Rattus norvegicus hyaluronan receptor for endocytosis HARE
DEFINITION precursor, mRNA, partial cds.
ACCESSION AY007370
VERSION AY007370.1 GI:24285892
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 4706)
AUTHORS Zhou, B., Weigel, J. A., Saxena, A. and Weigel, P. H.
TITLE Molecular Cloning and Functional Expression of the Rat 175-kDa Hyaluronan Receptor for Endocytosis
JOURNAL Mol. Biol. Cell 13 (8), 2853-2868 (2002)
MEDLINE 22169209
PUBMED 12181351

REFERENCE 2 (bases 1 to 4706)
AUTHORS Zhou, B., Weigel, J. A. and Weigel, P. H.
TITLE Direct Submission
JOURNAL Submitted (25-AUG-2000) Department of Biochemistry and Molecular Biology, University of Oklahoma Health Sciences Center, 940 Stanton L. Young, BMSB 860, Oklahoma City, OK 73104, USA

FEATURES
source
1..4706
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
<1..4296
/codon_start=1
/product="hyaluronan receptor for endocytosis HARE
precursor"
protein_id="AAG13634.1"
/db_xref="GI:24285893"
/translation="SLESLTRLEQMPDYSIFRGYIHYNLAIESADAYTVFVFN
EAIENYREKATSLKEDILRYHVVLGKLLKNDLHNGMRETMFGSYLLAFRLND
OLYNEAPINVTNATDKVHGLEKLEIKRCDNDNTIIVRGECKSQQAPCP
ETKPLRETRKCLISLYFMGKRSVFIGQPCQVRLITRACCAAGFPQCAQCGRGN
VCSGDCVGRGVKCDMEITTDNCGTCTSANCLDPDGKASCKAGFRNGTCT
AINACTSNCGSCSTKADKRTTGNRVCKAGTGDGIVCLEINPCLNHHGGCDRA
ECTGTGNQAVCNCLPKYTGKGVCSLINVCLTNNGCSPFAFCNCTEQQRICTKCP
DYTDGIVCGRSIVGELPKNPSTQYFQLEHVAHRELAPGPTVPAFISFNHFP
RIKWDQGLMSQVLRVHVVGCCQLLDLITPVSATLLOGEPVSI SVSDTVPINNE
AKVLSDIISTNGVHVHIDKLSPLNLTIPKDALGRVLQNLTVAAHNGYTKSKLI
QDGLSVITDSITHPTVFWPTKALEALPPEQDFLNQDNKDKKSLKLFVIR
SKALASLPRASWKTQSGELSVKGTGSDIGELFNEQCRFIHRLGLFDVGVAG
IDCLMNPITLGRCDTFTTFDIPOECGSCITPKPKLSPKPKVKKCIYNPLPFRN
VEGQNLCTTVIQTFRCHGYFMPDCAQCPGPDPTPCNNRGMCRDLVTPMQCLCHTG

CDS
1..4296
/product="hyaluronan receptor for endocytosis HARE
precursor"
protein_id="AAG13634.1"
/db_xref="GI:24285893"
/translation="SLESLTRLEQMPDYSIFRGYIHYNLAIESADAYTVFVFN
EAIENYREKATSLKEDILRYHVVLGKLLKNDLHNGMRETMFGSYLLAFRLND
OLYNEAPINVTNATDKVHGLEKLEIKRCDNDNTIIVRGECKSQQAPCP
ETKPLRETRKCLISLYFMGKRSVFIGQPCQVRLITRACCAAGFPQCAQCGRGN
VCSGDCVGRGVKCDMEITTDNCGTCTSANCLDPDGKASCKAGFRNGTCT
AINACTSNCGSCSTKADKRTTGNRVCKAGTGDGIVCLEINPCLNHHGGCDRA
ECTGTGNQAVCNCLPKYTGKGVCSLINVCLTNNGCSPFAFCNCTEQQRICTKCP
DYTDGIVCGRSIVGELPKNPSTQYFQLEHVAHRELAPGPTVPAFISFNHFP
RIKWDQGLMSQVLRVHVVGCCQLLDLITPVSATLLOGEPVSI SVSDTVPINNE
AKVLSDIISTNGVHVHIDKLSPLNLTIPKDALGRVLQNLTVAAHNGYTKSKLI
QDGLSVITDSITHPTVFWPTKALEALPPEQDFLNQDNKDKKSLKLFVIR
SKALASLPRASWKTQSGELSVKGTGSDIGELFNEQCRFIHRLGLFDVGVAG
IDCLMNPITLGRCDTFTTFDIPOECGSCITPKPKLSPKPKVKKCIYNPLPFRN
VEGQNLCTTVIQTFRCHGYFMPDCAQCPGPDPTPCNNRGMCRDLVTPMQCLCHTG

FNGTACELCHWGRFGPDOPRSCSBHQDDEGITSGECLCETGWTATASCDTPTAVFA
 VCTPACSVHATCENNTVCNLYEGDITCIVDFCKQNNGGCAKAKSKGQKQTVS
 CSCKGKGDGYSCEIIEPCADGVNGGCHETACRTMGKHKCECKSHYVDGVDCE
 PEQLDLRCLQNGQCHPDSCADLYFOOTTVGVFHLRSPLOQYKLTDFDKAKEACAKE
 AATITATYNOLSYAKAKYHLKCGALMESGRVAYPTTYASOKGCANVGVIVDYSRANK
 SEMWDVFCYRMKDVNCTKAGYVGDGFCSGNLOVLMSFPLTNFLDEVLAFSKSA
 RGOAFLKLTDLISIRGLEVPONSLPGNKLSGRDIEHLLTNVNVSFNDLYNGTFL
 RYMLGSLLITVSQDLQHEFTFVDRSILQWDDITAAAGLILHIISEPPURAPPTATAA
 HTMLSGTIFCAVVITGALAAAYSIFRLKQRTTFQFDFQDKTLMMLASSPRIS
 QTLQWRPQRHPQSPVPTSPQTLNRIWRTATLWGHCGDPMRSQAATTVPTR"
 1. .4293
 mat_peptide
 /product="hyaluronan receptor for endocytosis HARE"
 /note="175 kDa; derived by proteolytic cleavage"

ORIGIN

Alignment Scores:

Pred. No.: 6,69e-139 Length: 4706
 Score: 1514.50 Matches: 280
 Percent Similarity: 87.54% Conservative: 29
 Best Local Similarity: 79.32% Mismatches: 43
 Query Match: 81.21% Indels: 1
 DB: 10 Gaps: 1

US-09-466-778B-11 (1-353) x AY007370 (1-4706)

QY	1	MetThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeu	20
Db	3022	ATGACGGGCCCCAGGCAAGATAGTGTGAATGTAAAGTCATATGTCCGGGACCGAGG	3081
QY	21	AsnCysGluProGluGlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis	40
Db	3082	GACTGTGACCTGAGCAGCTGGCGCTGACCGTGTCTACAGACACACGACAGTGCAC	3141
QY	41	AlaAspAlaLysCysValAspLysPheGlnAspThrThrValGlyValPheHisLeu	60
Db	3142	CCAGATGCCAGCTGTGCAGACCTCTACTTCCAGGACACGCGTAGGATATTCCTACTA	3201
QY	61	ArgSerProLeuGlyGlnTrpLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn	80
Db	3202	CCCTCCCACTGGCGCAGTACAACTGACATTTGACAAAGCAAGACCTGTGCCAAA	3261
QY	81	GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr**GlnLysAlaLysTyrHis	100
Db	3262	GAAGCTGGACCATAGACCACTACACAGCTCTCTATGCCAGACGACCAAGTATACAC	3321
QY	101	LeuCysSerAlaGlyTrpLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer	120
Db	3322	CTGTGCTGGCGCGCTGGCTGGAGAGTGGCGGGTTCCTACCCGACTACGTATGCTCT	3381
QY	121	GlnAsnCysGlySerGlyValValGlyLysValAspTyrGlyProArgProAsnLysSer	140
Db	3382	CAGAAGTGTGTGCAACAGCTGTGGATCGTAGACTAGATCCAGGCCCAACAGAGT	3441
QY	141	GluMetTrpAspValPheCysTyrArgMetLysAspValAsnCysThr**LysValGly	160
Db	3442	GAAATGTGGATGTCTCTGTGTACCGGATGAAGATGTGAATGCACTGCAAGCAGGC	3501
QY	161	TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuLeuGlnValLeuMetSerPhePro	180
Db	3502	TATGTGGGAGATGGCTTCTGTGGAGTGGGAACCTGCTGCAGGTCTCATGTCTCC	3561
QY	181	SerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerAlaArgGlyArg	200
Db	3562	TCGCTCACAACTTCTGCACAGGTGTGGCTTTTCCAAAGAGCTACCCCGAGACAG	3621
QY	201	AlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsn	220
Db	3622	GCATTTTGAACACCTGACTGACTGCTCATCTCCGTGGCACCTGTTGTGCCACAGAAC	3681
QY	221	SerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisLeuAlaAsn	240
Db	3682	AGTGGGCTACCGGGAAATAAGACCTGTCTGCCCGGACATTGAGCACCACTACTAAT	3741

QY	241	ValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGly	260
Db	3742	GTCAACGCTCTCTTTTACAATGACCTTCTCAATGTTACCTTCTGAGGACTATGCTGGGA	3801
QY	261	SerLysLeuLeuLeuThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal	280
Db	3802	AGCCAACTGCTCATTTACCTTCAGCCAGGACGAGCTCCAC--CAAGAGACCAAGTTTGTG	3858
QY	281	AspGlyArgAspThrLeuGluTrpAspIleCysAlaSerAsnGlyIleThrHisValIle	300
Db	3859	GATGGAAGATCCATCTTCGAGTGGGACATCATCGCGCAATGGAATCTCCATATTATT	3918
QY	301	SerArg**LeuLysAlaProProAlaProValThrLeu**HisThrGlyLeuGly***	320
Db	3919	TCTGAACCTTTGAGAGCTCTCCACAGCAGCAACGGCTGCCCACTCTGGCGCTGGGACA	3978
QY	321	GlyIlePhe*****IleIleLeuValThrGlyAlaValAlaLeuAlaAlaTyrSerTyr	340
Db	3979	GGTATATTCTGCGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	4038
QY	341	PheArgIleAsnArgLysThrIleGlyPhe**HisPhe 353	
Db	4039	TTCCGGCTAAAGCAGCAGCAACCACTGTTTCCAGCGTTT 4077	

RESULT 14

AF364951
 LOCUS AF364951 8157 bp mRNA linear ROD 25-MAR-2002
 DEFINITION Mus musculus stabilin-2 mRNA, complete cds.
 ACCESSION AF364951
 VERSION AF364951.2 GI:19705588
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 8157)
 AUTHORS Politz,O., Gratchev,A., McCourt,P.A.G., Schledzewski,K.,
 Guillot,P., Johansson,S., Svineng,G., Franke,P., Kamnitsch,C.,
 Kzyshkowska,J., Longati,P., Veltan,F.W., Johansson,S. and
 Goerdts,S.
 TITLE Stabilin-1 and -2 constitute a novel family of fasciclin-like
 hyaluronan receptor homologues
 JOURNAL Biochem. J. 362 (Pt 1), 155-164 (2002)
 MEDLINE 21818465
 PUBMED 11829752
 REFERENCE 2 (bases 1 to 8157)
 AUTHORS Guillot,P., Schledzewski,K., Gratchev,A. and Goerdts,S.
 TITLE Direct Submission
 JOURNAL Submitted (28-MAR-2001) Dermatology, Universitätsklinikum Mannheim,
 Universitat Heidelberg, Theodor Kutzer Ufer 1-3, Mannheim, BW
 68135, Germany
 3 (bases 1 to 8157)
 AUTHORS Guillot,P., Schledzewski,K., Gratchev,A. and Goerdts,S.
 TITLE Direct Submission
 JOURNAL Submitted (25-MAR-2002) Dermatology, Universitätsklinikum Mannheim,
 Universitat Heidelberg, Theodor Kutzer Ufer 1-3, Mannheim, BW
 68135, Germany

REMARK

sequence update by submitter

COMMENT

On Mar 25, 2002 this sequence version replaced gi:19568177.

FEATURES

Location/Qualifiers
 1..8157
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="BALB/c"
 /db_xref="taxon:10090"
 /tissue_type="liver"
 62..7741
 /note="endothelial protein"
 /codon_start=1
 /product="stabilin-2"
 /protein_id="AA191684.2"
 /db_xref="GI:19705588"
 /translation="MARSKLLLGKLLPLILFLGLLVQNASPTPEAFELTKRCDKST"

CDS

LTIKTBQCSVNTAVKCPDGYIKITNGTVGRDVRSLKLSQSVLDIPGRHICRKO
 YLQPCPCFHWGPMCEGPGARAPKCGRGVDCMEGTGSCSRAGRGRFACNCA
 DVFPGNCSAVCSVHNGSISGDTGCELSAYGRPCDKPIPECAALCPENSR
 SPSSKDETKLCKLPSYKGGQTKPIPNCLKNVCHPHASCSVLGNRHSVCQKGY
 QGDQVCLPCTVOTSTKSVRYDVGSHCEKHEYNFVPGVCSMTDIC
 ESKNPKHANCSTVSPGTOCTCKGVVGLNCGNIMORLRELNTPRGMWGOL
 TSFISILDRTYAPLNSLPTVLLPSDKGLKGVYDVKELLMKKAARFVYKHLIAG
 MSTOMYMLDTYTLTGQSGEITLNDKQNLKGLYKSIQVQIIGNIVASNGLVHL
 DRAMDKIETLESNQQTIMTLOPRYKSLLEKTNVQLEKGGDEPYTIFVPS
 NEALSMTAGVLDVLLSPGSKLLELVHYVAFVQLEVALTVLTHIRSMANQIIT
 FNISSKGOILANNVADTEVAANKGRIYTLTVLPSPILPILPHRCNTRKMKILG
 TVCRFMKNCPCPNSTPALTFTNKCFSGRANWKIGICARYCDVTPRCKGFF
 GPDNCPKPGFMPNCGQCIIDGLGGNGTCICEDFGSGFCQPCSKENRYGPOCNR
 CQVHIGICNRLSDGSLPTCIBRTAGRCFQKQTSACGPMFQFCHHATCESNET
 ASCVNDGEGDGLTSCDKKPCGLSTSGGSPNABCIQASTGYSCVQCRQGTNGR
 DCEVINSCLLPSGGSHDANATCLYVPOQNECECKGFRNGIDCEPILESLETK
 HPLATQVTLTSGVNSCVQEGVNGVLCYGNVLMELSFLEAAVYFQWVNNASIQSM
 LSATSNLVVPSLQAIKMDQNEKSPWLSRNNTPALIKYHTLLGTBYVALDLTPSS
 HMLATSLGSLRLDKADGNITIEGASVVDGNDATNGVHIINKVLIPORLTGSLLP
 SILTLREMPDYSIPRGYIHYNLASALEADATVTVFVNNNEATSYIREKKATSLKE
 DILQHVVLGKRLNDLHNGHRETMIGFSYLLAFPLHNDOLYVNEAPINTVATD
 KGVTHGLSKVLEIKKRNNDNTIIVRGCKSQOCLPLETLPSETRKCIYVYF
 MGRSIFITGCOLQCVRTIITISACCAFPQPCQACPCGKQVCSNGFCLDGVNGTGT
 CECEGNGTACETCEKGYIHCDOACSCHVRCNOQPSGDCDCDVGWRGVKCD
 EITDNCNGTCHTSANCLLDPGKASCKAAGFGNGTVCTAINACETISNGGCSAKAD
 CKRTIPSRVCVKAGYGDGIVCLINPCLNMGCDRHAECTGTGPNVAVCNCLPK
 YTGDKVCTLLNVLCTNNGGSPFAPCNHTEQDQRTCTCKEDYTGDGIVCRGSHSEL
 PKNPSTSYFQLOHBOELAGPFTFVPSSDSFNSKLVKVDKQGLMSQLIRY
 HVACQQLLENLKITTSATLQGEPIISVSNQDVLTKKAKVLSIIITNGVIVH
 DITLSPONLLITPKAGSRVLLNTTVAANHGYTKFSKLQDSLLKVIITDPMTPV
 TLFWPTDKALQALPOEDFLNEDNKKLAKYLFHVIRDTMALASDLPRASWKT
 QGSELVRCGSDVGEFLNQMCRILQRLIFPGGVAYGIDCLLMDPTEGRCDF
 TITNIPGSCGSDVGEFLNQMCRILQRLIFPGGVAYGIDCLLMDPTEGRCDF
 CSYPMPCQACQGPDPFCNNRNGCYDQYKPTGQCQCHTFNGFACSLCLPGRGPD
 CQPCSHGQCEGIDTSGOCLCEAGTFRCDAPTIVIPVPCISPMHATCMENNT
 VCNVNLBGDGTCTVDFPCKNNGCAKVAKCSOKGVQVSCQKGVKGDHSGTEI
 DPCANGVYGECHETARTGTPGKQKCECKSHYVGDGDCBEPOLPLRCLQDNGQCH
 PDANCVDLHQDTTVGVFHLRSLPGLQYKLTIDKAKCAKAAATATYNLSYAKAK
 YHLCAGLSEGRVAYPTIYASKCANIVGIDVYTRNKSSEMDFVFCRMRKDVNCTC
 KAGYVGDGFCNGLLQVLMSPSLNFTLTVLVSRAQGRAPLHLTDLISGTL
 FVPONSGLPKNSGLQVLEHLLTVNVSFYVDLTVLKLRLGSLQIITSSQDLH
 OEARVDCRALOWDIIASNGVHLIISEPLKAPPTAATAHSGLTGFCFVAVLVTA
 IALAAYSERLNRORTGPRPESEDDIADALAFKQKQPSINPLNYETSTPAAPBPSCD
 PPTDSGERELNSDFLGALRS"

ORIGIN

Alignment Scores:
 Pred. No.: 6,89e-136 Length: 8157
 Score: 1487.00 Matches: 275
 Percent Similarity: 86.97% Conservative: 32
 Best Local Similarity: 77.90% Mismatches: 44
 Query Match: 79.73% Indels: 2
 DB: 10 Gaps: 2

US-09-466-778B-11 (1-353) x AF364951 (1-8157)

QY 1 MetThrGlyProGlyHisHisCysGluCysLysSerHisTyrValGlyAspGlyLeu 20
 Db 6512 ATGACGGGGCCAGGCAAGAGTGTGAATGTAAAGTCACTATGTCCGGGATGGAAG 6571
 QY 21 AsnCysGluProGluGlnLeuProIleAspArgCysLeuGlnAspGlyGlnCysHis 40
 Db 6572 GACTGTGAGCCCAACAGCTGCCCCCTTGACCGCTGTTCAGGACCAACGCAATGCCAC 6631
 QY 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60
 Db 6632 CCAGATGCCAACTGTGTGGACCTTCATTCCAGGACACTACTGTTCGGCGTATTCACCTA 6691
 QY 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80
 Db 6692 CGTCTCCCACTGGGCCAGTACAACTGACATTTGACAAAAGCCAGGCGCTGTGTGTAAG 6751
 QY 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr**GlnLysAlaLysTyrHis 100

Db 6752 GAAGCTGATCCATAGCACCTCAACACAGCTCTCCCTATGCCAGAGCCAGATATCAC 6811
 QY 101 LeuCysSerAlaGlyTyrLeuGluThrGlyArgValAlaTyrProThrAlaPheAlaSer 120
 Db 6812 CTCTGCTCAGCTGCTGGCTGGAGAGTGGGGGTTCCTTACCCACGATCTACGCTCT 6871
 QY 121 GlnAsnCysGlySerGlyValValGlyIleValAspTyrGlyProArgProAsnLysSer 140
 Db 6872 AAGAAATGT---GCAAAACATTTGCGGATTTGTAGACTACGGAACAGGACCAACAAGAGT 6928
 QY 141 GluMetTyrAspValPheCysTyrArgMetLysAspValAsnCysThr**LysValGly 160
 Db 6929 GAATGTGGAGCGTCTTCTGTTCAGATGAAAGATGTTACTGCACCTGCAGGACGAGC 6988
 QY 161 TyrValGlyAspGlyPheSerTyrSerGlyAsnLeuLeuGlnValLeuMetSerPhePro 180
 Db 6989 TACGTGGCGGATGGCTTCTCTGCAATGGGAACCTGCTGCAAGTCTCATGCTCTTCC 7048
 QY 181 SerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyArg 200
 Db 7049 TCCTCAGCAATCTCTGACAGAGGTGCTGTTTTTCCAGGAGCTCAGCCCAAGGCGG 7108
 QY 201 AlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAsn 220
 Db 7109 CGCTTTTGAACACCTTGACTGCTGCTCATCAGTGGCACCTGTTGTGCCACAGAAC 7168
 QY 221 SerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisHisLeuAlaAsn 240
 Db 7169 AGTGGGCTACCGAAAAATAAGAGCTGTCTGGGGGACATTGAGCACCACTCACTAAT 7228
 QY 241 ValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuGly 260
 Db 7229 GTCAACGCTCTCTTACGATGACCTGTCAATGATGATCCCTGAGGAGCTAGGCTGGGA 7288
 QY 261 SerLysLeuLeuLeuThrAspArgGlnAspProLeuHisProThrGluThrArgCysVal 280
 Db 7289 AGCCAACTGCTCATCCTCCAGCAGGACCCAGCTCCAC---CAAGAGCCAGGTTCGTG 7345
 QY 281 AspGlyArgAspThrLeuGluTyrAspIleCysAlaSerAsnGlyIleThrHisValle 300
 Db 7346 GATGGAAGAGCCATTCTGCAGTGGGACATCATTTGCTCTAAAGGGGTTCCTCATATCAT 7405
 QY 301 SerArg**LeuLysAlaProAlaProValThrLeu**HisThrGlyLeuGly*** 320
 Db 7406 TCTGAACCTTTGAAGAGTCTCTCCAGCGCCGACCGCTGCCACTCTGGCCCTAGGAACA 7465
 QY 321 GlyIlePhe****IleIleLeuValThrGlyAlaValAlaLeuAlaAlaTyrSerTyr 340
 Db 7466 GGCATATTCTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7525
 QY 341 PheArgIleAsnArgLysThrIleGlyPhe***HisPhe 353
 Db 7526 TTCGGCTTAAACGAGAGAACCACTGGCTTCCGGCGTCTT 7564

RESULT 15

AX149447
 LOCUS AX149447
 DEFINITION Sequence 3 from Patent WO0136638.
 ACCESSION AX149447
 VERSION AX149447.1 GI:14347929
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Shinkets R.A., Lichenstein H., Vernet C. and Fernandes E.
 TITLE Polypeptides and nucleic acids encoding same
 JOURNAL Patent: WO 0136638-A 3 25-MAY-2001;
 FEATURES
 Location/Qualifiers
 source 1..2011

Job time : 4966 secs

/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores: 1.16e-108 Length: 2011
Pred. No.: 1203.00 Matches: 230
Score: 1203.00
Percent Similarity: 84.87% Conservative: 0
Best Local Similarity: 84.87% Mismatches: 4
Query Match: 64.50% Indels: 37
DB: 6 Gaps: 1

US-09-466-778B-11 (1-353) x AX149447 (1-2011)

```
QY 1 MetThrGlyProGlyLysHisLysCysGluCysLysSerHisTyrValGlyAspGlyLeu 20
DB 1309 ATGACAGGCCCGGCGAGCAAGGTGAGTGAAGTCACTATGTCGGAGATGGGCTG 1368
QY 21 AsnCysGluProGluGlnLeuProIleAspArgCysLeuGlnAspAsnGlyGlnCysHis 40
DB 1369 AACTGTGAGCGGAGCAGCTGCCATTGACCGCTGCTTACAGGACAATGGCGAGTGCAT 1428
QY 41 AlaAspAlaLysCysValAspLeuHisPheGlnAspThrThrValGlyValPheHisLeu 60
DB 1429 GCAGAGGCCAAATGTCCGACCTCCACTCCAGGATACCACTGTGGGTGTTCCATCTA 1488
QY 61 ArgSerProLeuGlyGlnTyrLysLeuThrPheAspLysAlaArgGluAlaCysAlaAsn 80
DB 1489 CGCTCCCGACCTGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCGCTGTGCCAAC 1548
QY 81 GluAlaAlaThrMetAlaThrTyrAsnGlnLeuSerTyr***GlnLysAlaLysTyrHis 100
DB 1549 GAAGCTGGACCATGGCAACCTCAACACAGCTCTCTATGCCCAAGGCCCAAGTACCAC 1608
QY 101 LeuCysSerAlaGlyTyrLeuGlnThrGlyArgValAlaTyrProThrAlaPheAlaSer 120
DB 1609 CTGTGCTCAGCAGGCTGGCTGGAGACGGGGGGTGGCTACCCACACAGCCTTCGCCCTCC 1668
QY 121 GlnAsnCysGlySerGlyValValGlyIleValAspTyrGlyProArgProAsnLysSer 140
DB 1669 CAGAACTGTGGCTCTGGTGTGGTGGATAGTGAGCTATGGACCCAGACCCCAACAAGAGT 1728
QY 141 GluMetTyrAspValPheCysTyrArgMet-LysAspValAsnCysThr***LysValG1 160
DB 1729 GAAATGTGGGATGTCTTCTGCTATCGGATGAAG----- 1762
QY 160 yTyrValGlyAspGlyPheSerTyrSerGlyAsnLeuLeuGlnValLeuMetSerPhePr 180
DB 1762 ----- 1762
QY 180 oSerLeuThrAsnPheLeuThrGluValLeuAlaTyrSerAsnSerSerAlaArgGlyAr 200
DB 1763 -----GAAAGTCTGGCCTATTCCAAACAGCTCAGCTCGAGGGCG 1800
QY 200 gAlaPheLeuGluHisLeuThrAspLeuSerIleArgGlyThrLeuPheValProGlnAs 220
DB 1801 TGCAATTTCTAGAACACCTCGACTGCTTCATCCGGCGCACCTCTTTGTGCCACAGAA 1860
QY 220 nSerGlyLeuGlyGluAsnGluThrLeuSerGlyArgAspIleGluHisLeuAlaAs 240
DB 1861 CAGTGGGCTGGGGGAGATGAGACCTTGTCTGGGGGGGACATCGNACACCACTCGCCAA 1920
QY 240 nValSerMetPhePheTyrAsnAspLeuValAsnGlyThrThrLeuGlnThrArgLeuG1 260
DB 1921 TGTACAGCATGTTTTCTACAATGACCTTGTCAATGGCACCACTGCAACCTGCAACGAGGCTGG 1980
QY 260 ySerLysLeuLeuIleThrAspArgGlnAsp 270
DB 1981 AAGCAAGCTGCTCATCACTGCCAGCCAGGAC 2011
```

Search completed: April 1, 2004, 11:22:00